

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using BW model

Run on: March 25, 2006, 16:11:03 ; Search time 3009.5 Seconds

(without alignments)
10652.835 Million cell updates/sec

Title: SBQ1-69C

Perfect score: 564
Sequence: 1 gcagccgcgcctcagaaac.....agtgcgcttcgcgacttc 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_by:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	562.4	99.7	10777	8 HS11Q13RP	Y12377 H. sapiens F
2	553	98.0	150110	14 AC097722	AC097722 Homo sapi
3	553	98.0	150214	8 AP006345	AP006345 Homo sapi
4	44.2	7.8	125020	8 AF429315	AF429315 Homo sapi
5	42.6	7.6	125020	8 AF429315	AF429315 Homo sapi
6	40.2	7.1	184533	8 AC093592	AC093592 Homo sapi
7	40.2	7.1	192573	8 AC073344	AC073344 Homo sapi
8	40.2	7.1	320658	14 AC145527	AC145527 Ateles
9	40	7.1	153129	9 AC163209	AC163209 Mus muscu
10	40	7.1	230041	9 AC111047	AC111047 Mus muscu
11	39.8	7.1	103034	14 AC067978	AC067978 Homo sapi
12	39.8	7.1	116039	8 AF281074	AF281074 Homo sapi
13	39.8	7.1	116039	8 AC007362	AC007362 Homo sapi
14	39.8	6.9	124551	14 AC159874	AC159874 Bos tauru
15	38.8	6.9	184532	9 AC124345	AC124345 Mus muscu
16	38.8	6.9	203106	9 AC021667	AC021667 Mus muscu
17	38.8	6.9	215604	14 AC160979	AC160979 Mus muscu
18	38.4	6.8	110000	14 CT005255_5	Continuation (6 of

C 19	38.2	6.8	2777	6 AX833454	AX833454 Sequence
C 20	38.2	6.8	2777	8 AK095212	AK095212 Homo sapi
C 21	38.2	6.8	74871	14 AK095212	AK095212 Homo sapi
C 22	38.2	6.8	178778	8 AC132936	AC132936 Homo sapi
C 23	38	6.7	110000	14 LMFCHR26_2	Continuation (3 of
C 24	38	6.7	119613	8 AC138582	AC138582 Pan trogl
C 25	38	6.7	121103	8 AP006623	AP006623 Homo sapi
C 26	38	6.7	142979	8 AC000029	AC000029 Homo sapi
C 27	38	6.7	190018	14 AC148588	AC148588 Pan trogl
C 28	38	6.7	205854	14 AC138186	AC138186 Homo sapi
C 29	37.8	6.7	111370	8 AC067815	AC067815 Homo sapi
C 30	37.8	6.7	134362	14 AC084853	AC084853 Homo sapi
C 31	37.8	6.7	138129	8 HS102D24	HS102D24 Human DNA
C 32	37.6	6.7	142979	8 AL355500	AL355500 Human DNA
C 33	37.4	6.6	184866	9 AL611934	AL611934 Mouse DNA
C 34	37.4	6.6	187911	14 AC136867	AC136867 Rattus no
C 35	37.4	6.6	207735	9 AC130279	AC130279 Mus muscu
C 36	37.4	6.6	213590	14 AC154101	AC154101 Mus muscu
C 37	37.4	6.6	223357	14 AC111958	AC111958 Rattus no
C 38	37.4	6.6	240039	14 AC111926	AC111926 Rattus no
C 39	37.2	6.6	21670	8 AF508041	AF508041 Homo sapi
C 40	37.2	6.6	36628	8 AF037222	AF037222 Human DNA
C 41	37.2	6.6	294817	14 AC114626	AC114626 Mus muscu
C 42	37	6.6	1422	4 BT020874	BT020874 Bos tauru
C 43	37	6.6	1785	8 HSTPRMPL3	HSTPRMPL3 Human chrom
C 44	37	6.6	2500	1 SSU65940	SSU65940 Streptomyce
C 45	37	6.6	157585	5 BX322664	BX322664 Zebrafish

ALIGNMENTS

RESULT 1
HS11Q13RP
LOCUS
DEFINITION
H. sapiens FGF-3 gene upstream flanking region.
ACCESSION
Y12377
VERSION
Y12377.1 GI:1934871
KEYWORDS
Alu repeat; FGF-3 gene; int-2 gene; L1 repeat; MIR repeat.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Djenabi,S., Brisson,O., Galdemard,C. and Lavialle,C.
Sequence analysis of the transcription control region upstream of the human FGF-3 gene
DNA Seq. 10 (4-5), 317-329 (1999)
JOURNAL
PUBMED
10727086
REFERENCE
2 (bases 1 to 10777)
Brisson,O.
Direct Submission
Submitted (04-APR-1997) O. Brisson, Lab. de Genetique Oncologique, CIRSURA1967, Institut Gustave-Rousey, 39 Rue Camille Desmoulins, 94805 Villejuif CEDEX, FRANCE
COMMENT
The BamHI site at the 3' end of this sequence (nt 10772) corresponds to the BamHI site at the 5' end of the FGF-3/int-2 gene sequence (nt 1; X1445), as ascertained by sequencing through this site.
Related sequence: B04369.
location/Qualifiers
1..10777
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="q13"
/tissue_type="placenta"
/clone_id="lambda D3, Clontech cat#HL1067J lot 1221"
/germline
489..631
repeat_region
/rpt_family="MIR2"


```
misc_feature 4045..9798
/note="assembly_name:Contig26"
gap 9799..9898
/estimated_length=unknown
misc_feature 9899..14019
/note="assembly_name:Contig27"
gap 14020..14119
/estimated_length=unknown
misc_feature 14120..19392
/note="assembly_name:Contig28"
gap 19393..19492
/estimated_length=unknown
misc_feature 19493..24872
/note="assembly_name:Contig29"
gap 24873..24972
/estimated_length=unknown
misc_feature 24973..43358
/note="assembly_name:Contig30"
gap 43359..44058
/estimated_length=unknown
misc_feature 44059..68992
/note="assembly_name:Contig31"
gap 68993..69092
/estimated_length=unknown
misc_feature 69093..104669
/note="assembly_name:Contig32"
clone_end:T7
vector_side:left"
gap 104670..104769
/estimated_length=unknown
misc_feature 104770..150110
/note="assembly_name:Contig33"
clone_end:SP6
vector_side:right"
```

ORIGIN

```
Query Match 98.0%; Score 553; DB 14; Length 150110;
Best Local Similarity 99.8%; Pred. No. 1.7e-121;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCAGCCCTGCTCAGAAAACGAAAGACGACGACCTCAGGTGACTCACCCCATGTG 60
DB 50478 GCAGCCCTGCTCAGAAAACGAAAGACGACGACCTCAGGTGACTCACCCCATGTG 50419

QY 61 GCTGAGGCGAGGAGGCTCTGAGGAGGCGAGGCGAGGCGTCAAGTGGTGAACGCA 120
DB 50418 GCTGAGGCGAGGAGGCTCTGAGGAGGCGAGGCGAGGCGTCAAGTGGTGAACGCA 50359

QY 121 GGGGTCTTGCCATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 50358 GGGGTCTTGCCATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 50299

QY 181 CTGGTGCAGGCTCTGGAAGTCTGGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG 240
DB 50298 CTGGTGCAGGCTCTGGAAGTCTGGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG 50239

QY 241 GCCGGCCATGCCCCCAACGTTGCACTCAGTCCCTTTGAGAGGTTGGGGCTTCCAGTC 300
DB 50238 GCCGGCCATGCCCCCAACGTTGCACTCAGTCCCTTTGAGAGGTTGGGGCTTCCAGTC 50179

QY 301 ACAAGGTCCCATCCACGTAACGAGCCGAGGTGCTGCAAGAGTCCCTCGAGTCATGAA 360
DB 50178 ACAAGGTCCCATCCACGTAACGAGCCGAGGTGCTGCAAGAGTCCCTCGAGTCATGAA 50119

QY 361 CCAAGGAGGCTTGGGAAACCACTCTGAAGGAGCATGGCTTTGAATTTAGTGAAGGATGG 420
DB 50118 CCAAGGAGGCTTGGGAAACCACTCTGAAGGAGCATGGCTTTGAATTTAGTGAAGGATGG 50059

QY 421 GGGTGGGCTGGGCAAGGCAACAGCTCTGAGTCAAGAGCCAGAGGCAAGAGGCAAGAGGCTGCTCC 480
DB 50058 GGGTGGGCTGGGCAAGGCAACAGCTCTGAGTCAAGAGCCAGAGGCAAGAGGCAAGAGGCTGCTCC 49999

QY 481 A-GCAGTGCCTGCGGCTCTGCGATGAGTCTCTGAGCACTGAGAAAGAGCTGTAGA 539
```

```
DB 49998 AGGCATGCGCGCGGCTCTGCGATGAGTCTCTGAGCACTGAGAAAGAGCTGTAGA 49939
QY 540 GAGGCAGTGGCGGCTTTCCGACTTC 564
DB 49938 GAGGCAGTGGCGGCTTTCCGACTTC 49914
```

RESULT 3

```
AP006345/c 150214 bp DNA linear PRI 06-AUG-2005
LOCUS Homo sapiens genomic DNA, chromosome 11 clone:RP11-186D19, complete
DEFINITION sequence.
ACCESSION AP006345
VERSION AP006345.4 GI:71891788
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
```

REFERENCE

```
1 Hattori,M., Toyoda,A., Taylor,T.D., Kuroki,Y., Fujiyama,A.,
Totohi,Y. and Sakaki,Y.
Homo sapiens genomic DNA
PUBLISHED Only in Database (2003)
2 (bases 1 to 150214)
Hattori,M., Toyoda,A., Taylor,T.D., Kuroki,Y., Fujiyama,A.,
Totohi,Y. and Sakaki,Y.
Direct Submision
Submitted (30-APR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hnp.gsc.riken.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Aug 5, 2005 this sequence version replaced gi:48290861.
```

FEATURES

```
source 1..150214
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-186D19"
```

ORIGIN

```
Query Match 98.0%; Score 553; DB 8; Length 150214;
Best Local Similarity 99.8%; Pred. No. 1.7e-121;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCAGCCCTGCTCAGAAAACGAAAGACGACGACCTCAGGTGACTCACCCCATGTG 60
DB 85889 GCAGCCCTGCTCAGAAAACGAAAGACGACGACCTCAGGTGACTCACCCCATGTG 85830

QY 61 GCTGAGGCGAGGAGGCTCTGAGGAGGCGAGGCGAGGCGTCAAGTGGTGAACGCA 120
DB 85829 GCTGAGGCGAGGAGGCTCTGAGGAGGCGAGGCGAGGCGTCAAGTGGTGAACGCA 85770

QY 121 GGGGTCTTGCCATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 85769 GGGGTCTTGCCATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 85710

QY 181 CTGGTGCAGGCTCTGGAAGTCTGGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG 240
DB 85709 CTGGTGCAGGCTCTGGAAGTCTGGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG 85650

QY 241 GCCGGCCATGCCCCCAACGTTGCACTCAGTCCCTTTGAGAGGTTGGGGCTTCCAGTC 300
DB 85649 GCCGGCCATGCCCCCAACGTTGCACTCAGTCCCTTTGAGAGGTTGGGGCTTCCAGTC 85590

QY 301 ACAAGGTCCCATCCACGTAACGAGCCGAGGTGCTGCAAGAGTCCCTCGAGTCATGAA 360
DB 85589 ACAAGGTCCCATCCACGTAACGAGCCGAGGTGCTGCAAGAGTCCCTCGAGTCATGAA 85530
```

Oy	361	CCAAAGGAGGCTTGGGAAACACACTCGAAGGGCAGTGGCTTTGATTTATGTAGAGAGGTGG	420
Db	85529	CCAAAGGAGGCTTGGGAAACACACTCGAAGGGCAGTGGCTTTGATTTATGTAGAGAGGTGG	85470
Oy	421	GGCTGGGCTGGGCAAGGSCACCAAGTGTGAGTCAGAGCCAGAGGCAAGAGCTGTGCTCCC	480
Db	85469	GGCTGGGCTGGGCAAGGSCACCAAGTGTGAGTCAGAGCCAGAGGCAAGAGCTGTGCTCCC	85410
Oy	481	A-GCATTGCCCGCGGCTTTCGATGCAAGTCTTCTTGACCACCTGAGAAACAGCCTGTAGA	539
Db	85409	AGGCATGCGCCGCGGCTTTCGATGCAAGTCTTCTTGACCACCTGAGAAACAGCCTGTAGA	85350
Oy	540	GAGGCAGTGGGCTCTTTCGCACTTC	564
Db	85349	GAGGCAGTGGGCTCTTTCGCACTTC	85325

RESULT 4	AF429315/c	LOCUS	AF429315	125020 bp	DNA linear	PRI 18-JAN-2007
DEFINITION	Homoe sapiens junctophilin 3 (JPH3) gene, partial cds.					
ACCESSION	AF429315.1					
VERSION	GI:17646244					

SOURCE ORGANISM	Homo sapiens (human) Homo sapiens
-----------------	--------------------------------------

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE	1 (bases 1 to 120520)
AUTHORS	Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S., Ingeroll, Ashworth, R.G., Fleisher, A., Stevanin, G., Brie, A., Potter, N.T., Ross, C.A. and Mergolis, R.L.
TITLE	A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2
JOURNAL	Nat. Genet. 29 (4), 377-378 (2001)

PUBMED	11694876
REFERENCE	2 (bases 1 to 125020)
AUTHORS	Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE	Direct Submission
JOURNAL	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA

```

FEATURES      location/Qualifiers
source        1. .125020

```

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

/chromosome="16"
/map="16q24.3; between D6S520 and WI-12410"
/note="isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
repeat_region
complement(35581..35746)

```

```
gene      /rpt_type=cadem  
          /rpt_unit="ctg"  
          complement (<36507. .>36887)
```

```

/gene="JP3"
/note="synonym: JP3"
complement(<36507..>36887)

```

```

/gene="JPH3"
/product="junctional in 3"
complement(<36507..36687)
CDS

```

```
/gene="JPH3"  
/note="component of the junctional complex between plasma  
membrane and endoplasmic reticulum"
```

```
/codon_start=1
/product="junctophilin 3"
/protein_id="AAI40941.1"
```

db_xref=GI:17646245
/translation="MSSGGRRNFDDGSSYCGNEDGKAHGVCTPPKQGEYTGSGHPEVLGYTPPSGNTYQGTMAQCKRIGLSEKGVVYGEWTHFKKR YGVRENGATYEGTWSNGDGYGTETYSDG"

Query Match	7.8%;	Score 44.2;	DB 8;	Length 125020;
Best Local Similarity	12.0%;	Pred. No. 3.2;		
Matches	64;	Conservative 222;	Mismatches 244;	Indels 2;
				Gaps 1;

23 AAGAGCGAGCACTCACTGCTACTACCCCACTGTGCTGGAGCGAGGAGGCTCTCT 82

83 GAGGCAAGGGCCAGGGCAGCCCGTCAGGTGGGTGACGGCAGGGGCTCTTGCCATGATGGGCAC 142

17481 YGSTRSMKKKKKGYSKYSRGMMKKKKKTCYCMKKYKYRRTSMCWYYMKSWGKYRYYKCC 17422

17421 MKKKGGCTGVRGMSKSKSGYKSMRMRGSSYSTSCPKSKCWGSWMKCMKMYMSYYKKRRMR 1736

203 TGGAGTGCACCAATGTTTCCCATTAAGGAAAGTGTGGCCGGCCATGCCCCCCCAACGTT 262

17361 GSMSSKCMRGYAGRGCTYSSMMSTRIKRRCSCYSYSYKKGKGMKMGGMKMGKGSKYWSSM 1730

17301 KKKKSSWSKCYSTKSYKSGRRKSGWGHSTSKAKASSMRRAGSKCTYGGSSITWSNNRRNN 1724

323 GCCCAGGCGCTGCAGAAGT--CCCTCCAGTCATGTAAACCAAGGAGGCTTTGGGAAC 380
::: :
17241 RRGKTCGCMYRRSRAMNNGNAAAAAGCTTCCCANTNNGGGGAAAAAGGGCGSARACY 1718
::: :
17241

381 CACATCTGAAGGCATGCGTTTGATTAGTAGAGAAGGCTGGGCGCTGGGCAAGGCCA 440

441 CCAGGTCTGAGTCAGAGCCAGAGGCAAGCTGTGTCCTCCAGACTGCCCCGCCCTGTG 500
17181 KGRMSKSCYRSQTRRRCMKSKCRGSRGKSMGMTRGRSGGKTSYSAKSGRGRCYYCMGW 17122

17121 GKGGRCKMSRRKIMKYKSYRRKRKWMTCMKYCSNYMAMYCRSMCCMKCKSCGGCYSMG 1706

501 CGATGCAGTCTCCTCTGGCCACCTGAGAACAGCCTGTGAGAGAGGCCAATGGCGCT 552
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
17061 MSSSYSGKYSWGMKSYYMNRSSYYSKRSTSLAMRSSIKRGMGTGGRYKGGGSR 17010
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::

RESULT 5

AF429315		DNA	linear	PRI 18-JAN-2002
AF429315	125020 bp			
Homo sapiens junctophilin 3 (JPJ3) gene, partial cds.				

ACCESSION AF429315
 VERSION AF429315.1
 KEYWORDS GI:17646244

SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Rukhrovita: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

1 (bases 1 to 125020) Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

**UNPUBLISHED
AUTHORS**
Holmes, S. E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H. S.,
Ingersoll-Ashworth, R. G., Fleisher, A., Stevanin, G., Brice, A.,
Parker, N. T., Doss, C. A., and Marnett, L. J.

TITLE	DOI
A repeat expansion in the gene encoding junctional protein-3 is associated with Huntington disease-like 2	10.1006/hl.2001.3272

REFERENCE

Ullmann, C., 25 (7), 571-576 (2002)
Mac: Genet.
PUBMED
11694876
2 (bases 1 to 125020)
Ullmann, C., 25 (7), 571-576 (2002)
Mac: Genet.
PUBMED
11694876
2 (bases 1 to 125020)

Page C A and Mastroianni B T.

AUTHORS ROBERT L. ANGELBELL, A.B., M.D., AND MAGGALIE A. B. HOLMES, S.B.
TITLE Direct Submission
JOURNAL *Psychiatry*, Johns Hopkins Medical
Institute, 600 N. Wolfe St. Baltimore, MD 21205
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institute, 600 N. Wolfe St. Baltimore, MD 21205

FEATURES	
source	IMBlications, 800 N. MOILE ST., BALTIMORE, MD 21207, COM
	Location/Qualifiers
	1. 125020
	1. revised items discarded

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"

```


	repeat_region	/map=16q24.3; between DIS520 and WI-12410" /note="isolated from a patient with Huntington's Disease-Like 2 (HDL2)" complement(13581..35746) /rpt_type=tandem /rpt_unit="ctg" complement(<36507..>36887) /gene="JPH3"
gene		/note="synonym: JP3" complement(<36507..>36887) /gene="JPH3" /product="junctionphillin 3" complement(<36507..>36887) /gene="JPH3"
mRNA		/note="component of the junctional complex between plasma membrane and endoplasmic reticulum" /codon_start=1 /product="junctionphillin 3" /protein_id="P4A0941.1" /db_xref="GI:17646245" /translacion="MSGGKFRNFDDGSSYCCGGNEDKAHGHCCTGPCKGGYYTGSWS HGEFLGVATVWPSGNTYQGTWAQGRKHGIGLESKGMVYKGEMTHGFKKRYGVRCAAG NGAYEGTWNSNGLDGYETYSBDG"
CDs		
ORIGIN		
Query Match	7.6%; Score 42.6;	DB 8; Length 125020;
Beeb Local Similarity	11.5%; Pred. No. 7.6;	
Matches	45; Conservative 170; Mismatches 175; Indels 0; Gaps 0;	
Oy	32 GCACACACCGGTGACTACCCCCCATGTGCTGGAAGCAGGAGGCCTCTGAGCGAGG 91	
Db	51746 KSAISCMYMSACKSSTCAKRSGCYKKTWGSMTYSTSRSMYTSTSYCTMWMWMSY 51805	
Oy	92 CCAGGCGACGCCGTGACGTGGGTGTCAGCGACAGGGCTCTTGCCATGCTGGACAGGGCTTC 151	
Db	51806 MSWTSKCYRTYSCKRKMCMCSMCYRSARCTWRSMWTSCTWYSRSTGMRSKMGWCXY 51855	
Oy	152 ATACAGCTTACTCAGTACAATGAGTCCCTGTGTCACAGCTCTGGAAGTCTGGAAGTGA 211	
Db	51866 RRCGCMWKTYCGKSKSWMKCATTTGMSCARKSMSCYGKICRPASCMGCVKCMSMT 51925	
Oy	212 GCATGTTTTCCCATTAAGAAAGTGTGTGCGCCGCGCATGCCCCCACATTGACACTCA 271	
Db	51926 CMMRSYTTSCCTCYGTCYCMRKRGYRRRSSYKSCYKKSMTXSMGMSCYRGSRYMSY 51985	
Oy	272 CTGCTTTGAGAGGTTGGGGCTTCCAGTCAAGAGTCCCATCAAGTACCAGCCAGGTG 331	
Db	51986 RYRMRKCMCMWTRTKSGKSYNSSMWOMMGKITSYMCASSYMSCRGMSATSMWGCC 52045	
Oy	332 GTGCGAGAAGTCCCTTCGACAGTCAATAAACAAAGGAGCTTGGAAACACATCTGAAG 391	
Db	52046 YMSYRSASMCWGSTGSMRSYKSTWRSWCSMRGSMYSBMRMGCGCASMSCYRKSRK 52105	
Oy	392 GGCAATGGCTTTGATTTAGTAGAGGAGGTGG 421	
Db	52106 RGSWKGSMYTKSKXTGKGMWGRGSGWSK 52135	
RESULT 6		
LOCUS	AC093592/c	
DEFINITION	Homo sapiens BAC clone RP11-50L21 from 4, complete sequence.	
ACCESSION	AC093592	
VERSION	AC093592.3 GI:16756323	
KEYWORDS	HTG.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.	
AUTHORS	Meyer,R., Doeber,A. and Bielicki,L.	
TITLE	The sequence of Homo sapiens BAC clone RP11-50L21	

JOURNAL Unpublished (2001)
 REFERENCE 2 (bases 1 to 184533)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 3 (bases 1 to 184533)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (07-NOV-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 184533)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 184533)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA (bases 1 to 184533)
 REFERENCE 6 (bases 1 to 184533)
 AUTHORS Wilson,R.K.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAY-2005) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 COMMENT On Nov 7, 2001 this sequence version replaced gi:15687339.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0050121

NOTICE:
 This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
 The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Moon,P.Y., Zhao,B., Frengen,E., Tateo,M., Catenease,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.rsgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pRACE3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is RP11-628B17. Actual start of this clone is at base position 1 of RP11-50L21; actual end is at base position 184533 of RP11-50L21.

Data from AC073344 and AC016032 was used to finish this clone. AC093592. Polymorphisms have been identified between AC073344 and

FEATURES
source
AC093592. Location/Qualifiers
1..184533
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-50L21"
/clone_1fb="RPC1-11"
/complement(19506..26531)
/gene="K1AA1430"
complement(join(19506..20820,26266..26531))
/gene="K1AA1430"
complement(join(20779..20820,26266..26531))
/gene="K1AA1430"
/note="Homo sapiens K1AA1430 protein, mRNA (CDNA clone
MGC:40366 IMAGE:5240222), complete cds.; H_NH0050121.1
This gene was based on gi(20988136)
Continued from H_NH0628B17.1"
/codon_start=3
/product="unknown"
/protein_id="AAV40976.1"
/db_xref="GI:6392942"
/translation="FLQIDKGPQNHFDQSVAPGKNYSFTREVRQIDRENRLK
ELSRQAEKRGSKTIPRSADHPKLYHSLNRQKQRIERENLTISLCHGWSAVT"
54035..54810
/note="Cpg_island (%GC=74.2, o/e=0.97, #CpGs=110)"
59958..61369
/note="CpG_island (%GC=76.4, o/e=0.85, #CpGs=171)"
72793..73003
/note="CpG_island (%GC=58.8, o/e=1.06, #CpGs=17)"
90414..91446
/note="CpG_island (%GC=71.1, o/e=0.68, #CpGs=80)"
ORIGIN
Query Match 7.1%; Score 40.2; DB 8; Length 184533;
Best Local Similarity 60.6%; Pred. No. 27;
Matches 66; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 18 AACAGAGGAGCGACACCTCAGCGTGACTGACCCCATGTGGCGAGGCGAGGAGGC 77
Db 60170 AGCTCCCGGCGCGGCGGCGCCAGGAGGCCCACTGTTGTGACTGAAGCTTGCGCCAGC 60111
QY 78 CTCCTGAGGCGAGGCGCGAGGCGCGTCAAGTGGTGTGACGCGAGGCGGTC 126
Db 60110 CTGAGCGCGGCGCGGCGCGGCGCCCTGAAGTGCCTGTGCGGCGGCGCC 60062
RESULT 7
AC073344/c 192573 bp DNA linear PRI 16-APR-2002
LOCUS AC073344 Homo sapiens BAC clone RP11-628B17 from 4, complete sequence.
DEFINITION AC073344
AC073344
VERSION AC073344.7 GI:19848493
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 192573)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
PUBMED 9847074
REFERENCE
2 (bases 1 to 192573)
Harris,A., Meyer,R. and Nguyen,C.
The sequence of Homo sapiens BAC clone RP11-628B17
JOURNAL Unpublished (2001)
REFERENCE
3 (bases 1 to 192573)
Waterston,R.H.
JOURNAL Submitted (14-JUN-2000) Genome Sequencing Center, Washington

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (15-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 192573)
Waterston,R.H.
Direct Submission
Submitted (30-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 192573)
Waterston,R.
Direct Submission
Submitted (16-APR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 30, 2002 this sequence version replaced gi:18677646.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wuston.wustl.edu
----- Summary Statistics
Center project name: H_NH0628B17

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.
MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McHersom, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc
SOURCE INFORMATION:
The RPC1-11 human BAC library was made from the blood of one male
donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanesi,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RP11-714G18. Actual start of
this clone is at base position 1 of RP11-628B17; actual end is at
base position 192573 of RP11-628B17.
Polymorphisms exist between AC073344 and AC093592. Data from
AC093592 was used to finish AC073344.
FEATURES
source
Location/Qualifiers
1..192573
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="4"
/clone="RP11-628B17"

```
repeat_region      /clone_11b="RPCT-11"
                    381..445
misc_feature        /rpt_family="MER1_type"
                    742..1072
misc_feature        /note="match to EST BF244578 (NID:gil1158509) "
                    742..856
misc_feature        /note="match to EST AA164371 (NID:gi7141318) zp04a09.r1"
                    857..1424
misc_feature        /note="similar to EST BM511659 (NID:gi18682802) tjs1d11.x1"
                    902..1218
misc_feature        /note="similar to Homo sapiens EST H84319 (NID:gi10629950)
                    y97a10.s1"
misc_feature        1003..1811
misc_feature        /note="match to EST BG616006 (NID:gi3637377) "
                    1079..1331
misc_feature        /note="match to EST AA309603 (NID:gi1961983) "
                    1228..1583
repeat_region      /note="match to EST H84320 (NID:gi1062991) y97a11.r1"
                    1251..1286
misc_feature        /rpt_family="GA-rich"
                    1255..1811
misc_feature        /note="match to EST BI754577 (NID:gi5746155) "
                    1367..1811
misc_feature        /note="match to EST BF114916 (NID:gi10984392) "
                    1383..1731
misc_feature        /note="similar to Mus musculus EST BB651105
                    (NID:gi16485359) "
                    1466..1811
misc_feature        /note="similar to EST BM729997 (NID:gi19051330) tjs1d11.y1"
                    1526..1784
misc_feature        /note="match to EST AA384441 (NID:gi2036761) "
                    1628..1780
misc_feature        /note="match to EST BE093868 (NID:gi84744218) "
                    1670..1815
misc_feature        /note="similar to Homo sapiens EST BF172452
                    (NID:gi13438666) "
                    2149..2457
repeat_region      /rpt_family="Alu"
                    2255..2256
misc_feature        /note="match to EST H84320 (NID:gi1062991) y97a11.r1"
                    3136..3164
repeat_region      /rpt_family="AT-rich"
                    3330..3634
repeat_region      /rpt_family="Alu"
                    3635..3919
repeat_region      /rpt_family="L1"
                    3965..4273
repeat_region      /rpt_family="Alu"
                    4375..4510
repeat_region      /rpt_family="Alu"
                    4590..4803
repeat_region      /rpt_family="MER1_type"
                    4806..5101
repeat_region      /rpt_family="Alu"
                    5136..5290
repeat_region      /rpt_family="Alu"
                    5530..5609
repeat_region      /rpt_family="MIR"
                    5610..5726
repeat_region      /rpt_family="L1"
                    5811..6097
repeat_region      /rpt_family="Alu"
                    6519..6553
repeat_region      /rpt_family="tRNA-Gln-CAG"
                    6679..6790
repeat_region      /rpt_family="MER1_type"
                    6693..6698
misc_feature        /note="match to EST AW881803 (NID:gi8043813) "
                    6791..7081
repeat_region      /rpt_family="Alu"
                    7082..7219
repeat_region      /rpt_family="MER1_type"
                    7417..7722
```

```
repeat_region      /rpt_family="Alu"
                    8027..8317
repeat_region      /rpt_family="Alu"
                    8381..8759
repeat_region      /rpt_family="MER2_type"
                    8760..9071
repeat_region      /rpt_family="Alu"
                    9072..9106
repeat_region      /rpt_family="MER2_type"
                    9107..9403
repeat_region      /rpt_family="Alu"
                    9404..9689
repeat_region      /rpt_family="MER2_type"
                    9690..9986
repeat_region      /rpt_family="Alu"
                    9988..10289
repeat_region      /rpt_family="Alu"
                    10290..10841
repeat_region      /rpt_family="MER2_type"
                    10842..10956
repeat_region      /rpt_family="Alu"
                    10957..11251
repeat_region      /rpt_family="Alu"
                    11252..11398
repeat_region      /rpt_family="Alu"
                    11400..11640
repeat_region      /rpt_family="MER2_type"
                    11641..11942
repeat_region      /rpt_family="Alu"
                    11943..12383
repeat_region      /rpt_family="MER2_type"
```

Query Match 7.1%; Score 40.2; DB 8; Length 192573;
Best Local Similarity 60.6%; Pred.No.27;
Matches 66; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy 18 AACAGAGGACGCGACGACGCTGCTCAACCCCATGTGGCTGAGGCGAGGAGC 77
Db 20242 AGCTGCCGCGCGCGCCGACGCGAGGAGCCGACCGCTGTGTACTGAACCTGCCAGC 20183
Qy 78 CTCTGAGGACGCGCGCGCGCGCGCTGAGGTGGTGACGCGAGGCGTC 126
Db 20182 CTGACGCGCGCGCGCGCGCGCGCGCGCTGAGTGGTGTGCGGCGGCGCC 20134
```

RESULT 8
AC145527/c 320658 bp DNA linear HTG 19-JUN-2003
LOCUS Atelerix albiventris clone LB4-283G23, WORKING DRAFT SEQUENCE, 21
DEFINITION unordered pieces.
AC145527
VERSION AC145527.1 GI:32996768
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Atelerix albiventris (middle-African hedgehog)
ORGANISM Atelerix albiventris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Insectivora; Erimacidae;
Erimacinae; Atelerix.
REFERENCE
AUTHORS Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hoesseini,R.,
Peng,Z., Malinov,I. and Rubin,E.M.
TITLE Direct Submission
JOURNAL Unpublished 2 (bases 1 to 320658)
REFERENCE Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hoesseini,R.,
Peng,Z., Malinov,I. and Rubin,E.M.
AUTHORS Direct Submission
TITLE Submitted (19-JUN-2003) Genome Sciences, Lawrence Berkeley National
JOURNAL Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
COMMENT Draft Sequence Produced by Berkeley PGA
Web site: http://pga.lbl.gov
Center Code: PGABERK
Project Name: E024-283G23

Bac Clone Name: LB4-283G23

Additional information on comparative analysis and ordering are available at:

http://pga.lbl.gov/cgi-bin/search_cvs.cgi?type=newvalue=SRBPFI
Funding agent: Programs for Genomic Applications (NHLSI)
If library name is LB4 to LB4, please see website
for a description: http://www-gsd.lbl.gov/chem/BAC.html
These libraries are available through the BACPAC Resources Center:
http://www.chori.org/bacpac/libraryres.htm as LBNL-1 to LBNL-4.

Summary Statistics:

Sequencing vector: pIasmid, pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2325: contig of 2325 bp in length
2326 2425: gap of unknown length
2426 4760: contig of 2335 bp in length
4761 4860: gap of unknown length
4861 7844: contig of 2984 bp in length
7845 7944: gap of unknown length
7945 10691: contig of 2747 bp in length
10692 10791: gap of unknown length
10792 15758: contig of 4967 bp in length
15759 15858: gap of unknown length
15859 18758: contig of 2900 bp in length
18759 18858: gap of unknown length
18859 23515: contig of 4657 bp in length
23516 29133: contig of 5518 bp in length
29134 33870: contig of 4637 bp in length
33871 33970: gap of unknown length
33971 39516: contig of 5546 bp in length
39517 39616: gap of unknown length
39617 45912: contig of 6296 bp in length
45913 46012: gap of unknown length
46013 52370: contig of 6358 bp in length
52371 52470: gap of unknown length
52471 64704: contig of 12234 bp in length
64705 64804: gap of unknown length
64805 76264: contig of 11460 bp in length
76265 76364: gap of unknown length
76365 94758: contig of 18394 bp in length
94759 94858: gap of unknown length
94859 118713: contig of 23855 bp in length
118714 118813: gap of unknown length
118814 148021: contig of 29208 bp in length
148022 148121: gap of unknown length
148122 178939: contig of 30818 bp in length
178940 179039: gap of unknown length
179040 211754: contig of 32715 bp in length
211755 211854: gap of unknown length
211855 263200: contig of 51346 bp in length
263201 320658: contig of 57358 bp in length.
320659 320658: contig of 57358 bp in length.

FEATURES

source
1..320658 "Ateleleix albiiventris"
/organism="Ateleleix albiiventris"
/mol_type="genomic DNA"
/db_xref="taxon:9368"
/clone="LB4-283G23"
2326..2425
gap /estimated_length=unknown
4761..4860
gap /estimated_length=unknown

gap 7845..7944
gap /estimated_length=unknown
gap 10692..10791
gap /estimated_length=unknown
gap 15759..15858
gap /estimated_length=unknown
gap 18759..18858
gap /estimated_length=unknown
gap 23516..23615
gap /estimated_length=unknown
gap 29134..29233
gap /estimated_length=unknown
gap 33871..33970
gap /estimated_length=unknown
gap 39517..39616
gap /estimated_length=unknown
gap 45913..46012
gap /estimated_length=unknown
gap 52371..52470
gap /estimated_length=unknown
gap 64705..64804
gap /estimated_length=unknown
gap 76265..76364
gap /estimated_length=unknown
gap 94759..94858
gap /estimated_length=unknown
gap 118714..118813
gap /estimated_length=unknown
gap 148022..148121
gap /estimated_length=unknown
gap 178940..179039
gap /estimated_length=unknown
gap 211755..211854
gap /estimated_length=unknown
gap 263201..263300
gap /estimated_length=unknown

ORIGIN

Query Match 7.1%; Score 40.2; DB 14; Length 320658;
Best Local Similarity 58.0%; Pred. No. 26;
Matches 91; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

QY 1 GCAGCCTGCTCTCAAGAAAGAGAGACGACACTCAGCGTACTCACCCTCATG 60
DB 316431 GCACCCAGTTGCAACAAAGAGTGTGGCCAC--AGCAGACTCACCCTCTGG 316375
QY 61 GCTGAGGCGAGGAGCCTCTCTGAGGCGGCGGAGGCGGAGCCCTCAGGTGGTACGGCA 120
DB 316374 TGCTTTGGGCTGTGGCGCGCGAGGCTGTCTCCAGCAACCGGTGAGTGTGGGAC 316315
QY 121 GGGGCTTGGCCATGGTGGGCAAGGGGCTGCATACAG 157
DB 316314 TGGCTCCGCCATCAAGCGGCAAGTGGCTTATGCG 316278

RESULT 9

AC163209 153129 bp DNA linear ROD 03-AUG-2005
DEFINITION Mus musculus chromosome 3, clone RP24-11708, complete sequence.
ACCESSION AC163209
VERSION AC163209.2 GI:71725541
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 153129)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus chromosome 3, clone RP24-11708
JOURNAL Unpublished
2 (bases 1 to 153129)
REFERENCE Birren,B., Nusbaum,C., Lander,E., Aboueleil,A., Allen,N.,

Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorrie, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karacas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talama, J., Teefaye, S., Theodore, U., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
Journal
Submitted (07-JUN-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
3 (bases 1 to 153129)

AUTHORS
Britten, B., Nusbbaum, C., Lander, E., Abouelell, A., Allen, N., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorrie, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karacas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talama, J., Teefaye, S., Theodore, U., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
Journal
Submitted (03-AUG-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
On Aug 3, 2005 this sequence version replaced gi:67003668. All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

FEATURES
SOURCE
1. 153129
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone_1fb="RP24-11708"
/clone_1fb="RPCT-24 Male Mouse BAC"
2843. .2893
/rpt_family="AT_r1ch"
repeat_region
complement(2985. .3180)

repeat_region
5441. .5470
/rpt_family="B4"
repeat_region
5748. .5876
/rpt_family="B4"
repeat_region
6984. .7110
/rpt_family="MLTIC"
repeat_region
7462. .7801
/rpt_family="ORRD"
complement(9528. .9910)
/rpt_family="MT2B"
repeat_region
9937. .10085
/rpt_family="MER5B"
repeat_region
10825. .10982
/rpt_family="TG)n"
complement(11051. .11261)
/rpt_family="B4"
repeat_region
12012. .12033
/rpt_family="A)n"
repeat_region
12907. .13123
/rpt_family="MLTIC"
repeat_region
13433. .13812
/rpt_family="MTB"
complement(14278. .14503)
/rpt_family="L1"
repeat_region
14745. .15433
/rpt_family="L1_MM"
complement(15710. .15820)
/rpt_family="L1MA4A"
complement(15823. .16050)
/rpt_family="L1MA4A"
repeat_region
16057. .16122
/rpt_family="TG)n"
complement(16123. .16262)
/rpt_family="B3A"
repeat_region
16576. .16671
/rpt_family="TG)n"
repeat_region
1676. .16856
/rpt_family="TG)n"
repeat_region
16862. .16925
/rpt_family="TG)n"
repeat_region
16928. .17091
/rpt_family="TG)n"
complement(17311. .17610)
/rpt_family="MLTIA0"
complement(17617. .19453)
/rpt_family="L1_MM"
18133. .18163
/note="single clone coverage"
repeat_region
19454. .20704
/rpt_family="L1"
repeat_region
20758. .20824
/rpt_family="MIR"
complement(20852. .21030)
/rpt_family="MIR"
21915. .21942
/rpt_family="TTTTTG)n"
complement(21943. .22157)
/rpt_family="B3"
repeat_region
22301. .22397
/rpt_family="L2"
complement(22732. .22787)
/rpt_family="L1"
repeat_region
22786. .23404
/rpt_family="L1_MM"
complement(23405. .23879)
/rpt_family="MERVL_LTR"
23880. .24023
/rpt_family="L1_MM"
24051. .24205
/rpt_family="TCTCC)n"

	/map="3"
	/clone="RP23-176H24"
misc_feature	/clone_1b="RPC1-23 Female Mouse BAC"
	1. .21291
	/note="wgs end extension
repeat_region	clone end:Sp6_
	1144. .1280
repeat_region	/rpc_family="B1_MM"
	2418. .2629
repeat_region	/rpc_family="L2"
	2641. .2696
repeat_region	/rpc_family="MER91A"
	3023. .3043
repeat_region	/rpc_family="TCTG)n"
	3086. .3126
repeat_region	/rpc_family="(CA)n"
	3126. .3176
repeat_region	/rpc_family="(GA)n"
	3289. .3449
unare	/rpc_family="L2"
	3572. .3576
unare	/note"<30 qual SNGL region"
	complement(5063. .5383)
unare	/note="single clone coverage"
	5464. .5470
unare	/note"<30 qual SNGL region"
	5498. .5503
unare	/note"<30 qual SNGL region"
	5539. .5544
unare	/note"<30 qual SNGL region"
	5564. .5568
repeat_region	/note"<30 qual SNGL region"
	complement(5599. .6129)
repeat_region	/rpc_family="Lx2B"
	6272. .6332
repeat_region	/rpc_family="(GA)n"
	7652. .7678
repeat_region	/rpc_family="ORR1B1"
	complement(8072. .8141)
repeat_region	/rpc_family="MIR"
	8906. .9058
repeat_region	/rpc_family="B4A"
	9622. .9758
repeat_region	/rpc_family="(TCC)n"
	complement(10521. .10594)
repeat_region	/rpc_family="Zaphod"
	complement(12702. .12858)
repeat_region	/rpc_family="B3"
	complement(12871. .13230)
repeat_region	/rpc_family="MTD"
	13255. .13275
repeat_region	/rpc_family="AT_rich"
	13286. .14525
unare	/rpc_family="Lx4"
	14031. .14049
repeat_region	/note="single clone coverage"
	14535. .14580
repeat_region	/rpc_family="(CAAT)n"
	complement(14607. .14691)
repeat_region	/rpc_family="Lx9"
	14812. .14846
repeat_region	/rpc_family="(CAAAA)n"
	15203. .15348
repeat_region	/rpc_family="B3A"
	15349. .15391
repeat_region	/rpc_family="(CA)n"
	15392. .15437
repeat_region	/rpc_family="B3A"
	complement(15846. .15975)
repeat_region	/rpc_family="MTD"
	16107. .16376
repeat_region	/rpc_family="MTC"
	18724. .18891

```

repeat_region      /rpe_family="L1MB8"
                    19356. 19561
repeat_region      /rpe_family="B3"
                    19661. 19988
repeat_region      /rpe_family="AT_rich"
                    19990. 20035
repeat_region      /rpe_family="(TCTA)n"
                    20124. 20180
repeat_region      /rpe_family="AT_rich"
                    compLength(20343). 20479
repeat_region      /rpe_family="B1_MM"
                    compLength(20694). 21089
repeat_region      /rpe_family="M7A"
                    21292. 21297
misc_feature        /note="clone boundary
                    clone end:SP6
                    site:EcoR"
                    21395. 21487
repeat_region      /rpe_family="(CAGGG)n"
                    21517. 21672
repeat_region      /rpe_family="L1MB8"
                    compLength(21914). 28115
repeat_region      /rpe_family="URR1A"
                    30025. 30065
repeat_region      /rpe_family="(TG)n"
                    compLength(30077). 30214
repeat_region      /rpe_family="RSINE1"
                    30216. 30311
repeat_region      /rpe_family="CT_rich"
                    compLength(30413). 30611
repeat_region      /rpe_family="B4"
                    31033. 31131
repeat_region      /rpe_family="(CCA)n"
                    32917. 33316
repeat_region      /rpe_family="(TG)n"
                    33334. 33515
repeat_region      /rpe_family="(TG)n"
                    33712. 33732
repeat_region      /rpe_family="AT_rich"

```

Query Match	7.1%	Score 40;	DB 9;	Length 230041;
Best Local Similarity	52.4%;	Pred. No. 30;		
Matches 88;	Conservative 0;	Mismatches 80;	Indels 0;	Gaps 0;

QY	301	ACAGGGTTCATCATCACTACCAAGCCCAAGGTGCTGCAGAAAGTCCCTCCCACTCATGAA	360
Db	21304	ACAAAGTAGTGTGCTGGAGGCCAAGGACAGAGGCAAGGGGACAGGCGCAGTACAGAGGCCAAG	21363
QY	361	CCAAAGGAGGCTTGGGAAACCAATCTGAAAGGCATGCTTTGATTATTAGTAGAGGGTGG	420
Db	21364	GACAGGAGCCAGGGGTGAGACTGAGGTAGGAGGCCAAGGCGGGGCCAGACAGAGGCCAAG	21421
QY	421	GGCTGGGCTGGGCAAGGCCACCAAGTCTAGTCAAGCCAGAGGCCAAG	468
Db	21424	GGCGGGGCTCAAGCCAGAGGCCAAGGGCCAGGCCAAGGCCAAGGCCAAG	21471

RESULT 11							
LOCUS	AC067978						
DEFINITION		103034 bp	DNA	linear	HTG 03-JUL-2000		
ACCESSION	AC067978						
VERSION	AC067978.2						
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	1 (bases 1 to 103034)						
AUTHORS	Bitren,B., Linton,L., Nusbaum,C. and Lander,E.						
TITLE	Homo sapiens chromosome 2, clone CTD-2350K7 map 2, WORKING DRAFT						

JOURNAL
 2 (bases 1 to 103034)
 Unpublished
 Birtren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldin,J., Barna,N., Bastien,V., Bede,F.,
 Boguslavsky,L., Boutgalter,B., Brown,A., Burkett,G.,
 Campolano,A., Caetle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-pierre,N., Grant,G., Hages,B., Hearford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karakas,A.,
 Klein,U., Laroque,K., Lamasares,R., Landers,T., Lenczky,J.,
 Levine,R., Lien,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Melarim,J., Meneus,L., Mihova,T., Miranda,C., Mianga,Y., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisanu,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testafaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 3, 2000 this sequence version replaced gi:7656720.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L9518
 Center clone name: 2350_X.7

----- Summary Statistics
 Sequencing vector: M13/ M77815; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 95545 bases at least Q40
 Consensus quality: 99446 bases at least Q30
 Consensus quality: 100858 bases at least Q20
 Insert size: 98000; agarose-efp
 Insert size: 101634; sum-of-contigs
 Quality coverage: 4.5 in Q20 bases; agarose-efp
 Quality coverage: 4.4 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 1585: contig of 1585 bp in length
 1586
 1586: gap of 100 bp
 1586
 3150: contig of 1465 bp in length
 3151
 3250: gap of 100 bp
 3251
 4742: contig of 1492 bp in length
 4743
 4842: gap of 100 bp
 4843
 6347: contig of 1505 bp in length
 6348
 6447: gap of 100 bp
 6448
 8790: contig of 2343 bp in length
 8791
 8890: gap of 100 bp
 8891
 11684: contig of 2794 bp in length
 11685
 11784: gap of 100 bp
 11785
 15905: contig of 4121 bp in length
 15906
 16006: gap of 100 bp
 21395: contig of 5330 bp in length

FEATURES
 source
 1..103034
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="CTD-2350K7"
 /clone_1fb="CTD Human BAC"
 1..1585
 /note="assembly_fragment"
 1586..1685
 /estimated_length=100
 1686..3150
 /note="assembly_fragment"
 3151..3250
 /estimated_length=100
 3251..4742
 /note="assembly_fragment"
 4743..4842
 /estimated_length=100
 4843..6347
 /note="assembly_fragment"
 6348..6447
 /estimated_length=100
 6448..8790
 /note="assembly_fragment"
 8791..8890
 /estimated_length=100
 8891..11684
 /note="assembly_fragment"
 11685..11784
 /estimated_length=100
 11785..15905
 /note="assembly_fragment"
 15906..16005
 /estimated_length=100
 16006..21395
 /note="assembly_fragment"
 21396..21495
 /estimated_length=100
 21496..26064
 /note="assembly_fragment"
 26065..26164
 /estimated_length=100
 26165..334512
 /note="assembly_fragment"
 334913..35012
 /estimated_length=100
 35013..46466
 /note="assembly_fragment"
 46467..46566
 /estimated_length=100
 46567..56982
 /note="assembly_fragment"
 56983..57082
 /estimated_length=100
 57083..71309
 /estimated_length=100
 71310..86559
 /estimated_length=100
 86560..103034
 /estimated_length=100
 103034: contig of 16375 bp in length.

gap	71309.	.71408	/estimated_length=100
mlec_feature	71409..86559	/note="assembly_fragment"	
gap	86560..86659	/estimated_length=100	
mlec_feature	86660..103034	/note="assembly_fragment"	
ORIGIN			
Query Match	7.1%;	Score 39.8;	DB 14; Length 103034;
Best Local Similarity	64.8%;	Pred. No. 36;	
Matches 59;	Conservative 0;	Mismatches 32;	Indels 0; Gaps 0;
OY	4	GCCCTGCTCAGAAAACAGAAGACGCAGCACTACCGTACTACATCCCCATGTGGCT	63
Db	92387	GCCCTCCCTCGGAAMACGAGAAAGCTGGGGCACCCCTCGTGGGCTCACCATTTGGAA	93046
OY	64	GGAGCGCAGGAGCGCTCTGAGCGACGGCCA	94
Db	93047	TGATGAGAGGTTCCTCCCAAGGCGCTGGCCA	93077
RESULT 12			
AF281074/c			
LOCUS	AP281074	111934 bp	DNA linear PRI 21-DEC-2000
DEFINITION	Homo sapiens neuropilin 2 (NRP2) gene, complete cds, alternatively spliced.		
ACCESSION	AF281074		
VERSION	AF281074.1	GI:11934947	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 111934)		
AUTHORS	Rossignol,M., Gagnon,M.L. and Klagesbrun,M.		
TITLE	Genomic organization of human neuropilin-1 and neuropilin-2 genes: identification and distribution of splice variants and soluble isoforms		
JOURNAL	Genomics 70 (2), 211-222 (2000)		
PubMed	11112349		
REFERENCE	2 (bases 1 to 111934)		
AUTHORS	Rossignol,M., Gagnon,M.L. and Klagesbrun,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-JUN-2000) Surgical Research, Children's Hospital, 320 Longwood Ave, Boston, MA 02115, USA		
FEATURES	Location/Qualifiers		
source	1..111934		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	<3518..>111422		
gene	/gene="NRP2"		
mRNA	join(<3518..3590,17760..17937,36403..36584,42688..42918,43395..44150,46123..46292,48101..48256,60728..60872,63132..63761,65955..66099,69934..70050,73043..73183,83879..84141,85679..85775,86988..87023,108610..108660,111118..>111422)		
	/gene="NRP2"		
	/product="neuropilin-2a(22)"		
	join(<3518..3590,17760..17937,36403..36584,42688..42918,43395..44150,46123..46292,48101..48256,60728..60872,63132..63761,65955..66099,69934..70050,73043..73183,83879..84141,85679..85775,86988..87023,108610..108660,111118..>111422)		
	/gene="NRP2"		
	/product="neuropilin-2a(17)"		
	join(<3518..3590,17760..17937,36403..36584,42688..42918,43395..44150,46123..46292,48101..48256,60728..60872,63132..63761,65955..66099,69934..70050,73043..73183,83879..84141,85679..85775,86988..87023,108610..108660,		
CDS			

111118 . 111422)
/gene="NRP2"
/note="receptor for VEGF165 and semaphorins class3"
/codon_start=1
/product="neuropilin-2a(22)"
/protein_id="AA641898.1"
/db_xref="GI:11934949"
/translation="MDMPLRTVWFLALYRSRQVRGQDDPCGGRLSKNAGYITSPG
YLPQSPHONCEWLVAPENPOKIVLNFNPFIEIKHCKDFEIRKGDSEADLLG
KHCNINAPRITISSGMLYIKFTSDYARAGAFSLRYEIPFTSGDESKFTSNGTI
ESPPPEKRYPHNLDCTFTLLAKRMEIILQFLIPLEHDLQVPEGSGCKDWMIDLWG
IIPHGLGICGCTKPTSELRSSTGLSLTFHTMDAVAKDGSFAYVYVHPEPLENO
CNVLCMESSRIANEOSASTYSDRWTPPOOSRLHGDDNCTPRLDSNKELYOVIDR
FLTLHLALATQGLSRFTONGYVKSJKLEVTNEDMVMYTRHGKNGVFPANDATE
VLNKLHAPLRTVIRLPQTMHSGIALLEIFGGRVTDAPISKNMLSGLIADSOI
SASTOSTEALFSPSARLVYSRSGMPPRIIOAQPGEMQLQVLDGTYTKVGIIGAG
GDSITAEALAFPRKFVYSLNGKMEIIOOPRPOQPLFEGMNHVTPPIRFPDI
PAQVAVYPRMSFAGIGMRLLEYLGDMDTDSPTVETGLPVKSEETTPYRTEBAT
EGGNGCFEDDKXLQLPBSGNCNPDLEPSCMWYDHAQMLRTTMASSSPNDDT
DNATLRQSDOSRQEGQARLISPPVHLFSPVCMFQYQATGKRVLAQVREASQS
KLWLVIEDQSGWKHGRILIPSYDMEYOIVEGVIGKRGSELAIDIRISTVPLE
NCPMEIPAFGENPKVDIPIHERBEGVEIDIDEYEVDMSSSASTSGSGASPSTDEK
SMVLTLPIRLTIAMSSLGVLGATCAAGLILCTCSYSGLSRSRCTLLIENVNELYD
GLKKVMNHNOKCSSEA"

CD5
join(3518 . 3590,17760 . 17937,36403 . 36584,42688 . 42918,
43995 . 44150,46123 . 46292,48101 . 48256,60728 . 60872,
63412 . 63761,65955 . 66099,69934 . 70050,73043 . 73183,
83879 . 84141,85679 . 85775,86988 . 87008,108610 . 108660,
111118 . 111422)
/gene="NRP2"
/note="receptor for VEGF165 and semaphorins class3"
/codon_start=1
/product="neuropilin-2a(17)"
/protein_id="AA641897.1"
/db_xref="GI:11934948"
/translation="MDMPLRTVWFLALYRSRQVRGQDDPCGGRLSKNAGYITSPG
YLPQSPHONCEWLVAPENPOKIVLNFNPFIEIKHCKDFEIRKGDSEADLLG
KHCNINAPRITISSGMLYIKFTSDYARAGAFSLRYEIPFTSGDESKFTSNGTI
ESPPPEKRYPHNLDCTFTLLAKRMEIILQFLIPLEHDLQVPEGSGCKDWMIDLWG
IIPHGLGICGCTKPTSELRSSTGLSLTFHTMDAVAKDGSFAYVYVHPEPLENO
CNVLCMESSRIANEOSASTYSDRWTPPOOSRLHGDDNCTPRLDSNKELYOVIDR
FLTLHLALATQGLSRFTONGYVKSJKLEVTNEDMVMYTRHGKNGVFPANDATE
VLNKLHAPLRTVIRLPQTMHSGIALLEIFGGRVTDAPISKNMLSGLIADSOI
SASTOSTEALFSPSARLVYSRSGMPPRIIOAQPGEMQLQVLDGTYTKVGIIGAG
GDSITAEALAFPRKFVYSLNGKMEIIOOPRPOQPLFEGMNHVTPPIRFPDI
PAQVAVYPRMSFAGIGMRLLEYLGDMDTDSPTVETGLPVKSEETTPYRTEBAT
EGGNGCFEDDKXLQLPBSGNCNPDLEPSCMWYDHAQMLRTTMASSSPNDDT
DNATLRQSDOSRQEGQARLISPPVHLFSPVCMFQYQATGKRVLAQVREASQS
KLWLVIEDQSGWKHGRILIPSYDMEYOIVEGVIGKRGSELAIDIRISTVPLE
NCPMEIPAFGENPKVDIPIHERBEGVEIDIDEYEVDMSSSASTSGSGASPSTDEK
SMVLTLPIRLTIAMSSLGVLGATCAAGLILCTCSYSGLSRSRCTLLIENVNELYD
GLKKVMNHNOKCSSEA"

CD5
join(3518 . 3590,17760 . 17937,36403 . 36584,42688 . 42918,
43995 . 44150,46123 . 46292,48101 . 48256,60728 . 60872,
63412 . 63761,65955 . 66099,69934 . 70050,73043 . 73183,
83879 . 84141,85679 . 85775,86988 . 87008,108610 . 108660,
111118 . 111422)
/gene="NRP2"
/note="receptor for VEGF165 and semaphorins class3"
/codon_start=1
/product="neuropilin-2b(5)"
/protein_id="AA641900.1"
/db_xref="GI:11934949"
/translation="MDMPLRTVWFLALYRSRQVRGQDDPCGGRLSKNAGYITSPG
YLPQSPHONCEWLVAPENPOKIVLNFNPFIEIKHCKDFEIRKGDSEADLLG
KHCNINAPRITISSGMLYIKFTSDYARAGAFSLRYEIPFTSGDESKFTSNGTI
ESPPPEKRYPHNLDCTFTLLAKRMEIILQFLIPLEHDLQVPEGSGCKDWMIDLWG
IIPHGLGICGCTKPTSELRSSTGLSLTFHTMDAVAKDGSFAYVYVHPEPLENO
CNVLCMESSRIANEOSASTYSDRWTPPOOSRLHGDDNCTPRLDSNKELYOVIDR
FLTLHLALATQGLSRFTONGYVKSJKLEVTNEDMVMYTRHGKNGVFPANDATE
VLNKLHAPLRTVIRLPQTMHSGIALLEIFGGRVTDAPISKNMLSGLIADSOI
SASTOSTEALFSPSARLVYSRSGMPPRIIOAQPGEMQLQVLDGTYTKVGIIGAG
GDSITAEALAFPRKFVYSLNGKMEIIOOPRPOQPLFEGMNHVTPPIRFPDI
PAQVAVYPRMSFAGIGMRLLEYLGDMDTDSPTVETGLPVKSEETTPYRTEBAT
EGGNGCFEDDKXLQLPBSGNCNPDLEPSCMWYDHAQMLRTTMASSSPNDDT
DNATLRQSDOSRQEGQARLISPPVHLFSPVCMFQYQATGKRVLAQVREASQS
KLWLVIEDQSGWKHGRILIPSYDMEYOIVEGVIGKRGSELAIDIRISTVPLE
NCPMEIPAFGENPKVDIPIHERBEGVEIDIDEYEVDMSSSASTSGSGASPSTDEK
SMVLTLPIRLTIAMSSLGVLGATCAAGLILCTCSYSGLSRSRCTLLIENVNELYD
GLKKVMNHNOKCSSEA"

MRNA
join(<518 . 3590,17760 . 17937,36403 . 36584,42688 . 42918,
43995 . 44150,46123 . 46292,48101 . 48256,60728 . 60872,
63412 . 63761,65955 . 66099,69934 . 70050,73043 . 73183,
83879 . 84141,85679 . 85775,86988 . 87008,108610 . 108660,
111118 . 111422)
/gene="NRP2"
/product="neuropilin-2b(5)"
/protein_id="AA641900.1"
/db_xref="GI:11934949"
/translation="MDMPLRTVWFLALYRSRQVRGQDDPCGGRLSKNAGYITSPG
YLPQSPHONCEWLVAPENPOKIVLNFNPFIEIKHCKDFEIRKGDSEADLLG
KHCNINAPRITISSGMLYIKFTSDYARAGAFSLRYEIPFTSGDESKFTSNGTI
ESPPPEKRYPHNLDCTFTLLAKRMEIILQFLIPLEHDLQVPEGSGCKDWMIDLWG
IIPHGLGICGCTKPTSELRSSTGLSLTFHTMDAVAKDGSFAYVYVHPEPLENO
CNVLCMESSRIANEOSASTYSDRWTPPOOSRLHGDDNCTPRLDSNKELYOVIDR
FLTLHLALATQGLSRFTONGYVKSJKLEVTNEDMVMYTRHGKNGVFPANDATE
VLNKLHAPLRTVIRLPQTMHSGIALLEIFGGRVTDAPISKNMLSGLIADSOI
SASTOSTEALFSPSARLVYSRSGMPPRIIOAQPGEMQLQVLDGTYTKVGIIGAG
GDSITAEALAFPRKFVYSLNGKMEIIOOPRPOQPLFEGMNHVTPPIRFPDI
PAQVAVYPRMSFAGIGMRLLEYLGDMDTDSPTVETGLPVKSEETTPYRTEBAT
EGGNGCFEDDKXLQLPBSGNCNPDLEPSCMWYDHAQMLRTTMASSSPNDDT
DNATLRQSDOSRQEGQARLISPPVHLFSPVCMFQYQATGKRVLAQVREASQS
KLWLVIEDQSGWKHGRILIPSYDMEYOIVEGVIGKRGSELAIDIRISTVPLE
NCPMEIPAFGENPKVDIPIHERBEGVEIDIDEYEVDMSSSASTSGSGASPSTDEK
SMVLTLPIRLTIAMSSLGVLGATCAAGLILCTCSYSGLSRSRCTLLIENVNELYD
GLKKVMNHNOKCSSEA"

MRNA
join(<518 . 3590,17760 . 17937,36403 . 36584,42688 . 42918,
43995 . 44150,46123 . 46292,48101 . 48256,60728 . 60872,
63412 . 63761,65955 . 66099,69934 . 70050,73043 . 73183,
83879 . 84141,85679 . 85775,86988 . 87008,108610 . 108660,
111118 . 111422)
/gene="NRP2"
/product="neuropilin-2b(5)"
/protein_id="AA641900.1"
/db_xref="GI:11934949"
/translation="MDMPLRTVWFLALYRSRQVRGQDDPCGGRLSKNAGYITSPG
YLPQSPHONCEWLVAPENPOKIVLNFNPFIEIKHCKDFEIRKGDSEADLLG
KHCNINAPRITISSGMLYIKFTSDYARAGAFSLRYEIPFTSGDESKFTSNGTI
ESPPPEKRYPHNLDCTFTLLAKRMEIILQFLIPLEHDLQVPEGSGCKDWMIDLWG
IIPHGLGICGCTKPTSELRSSTGLSLTFHTMDAVAKDGSFAYVYVHPEPLENO
CNVLCMESSRIANEOSASTYSDRWTPPOOSRLHGDDNCTPRLDSNKELYOVIDR
FLTLHLALATQGLSRFTONGYVKSJKLEVTNEDMVMYTRHGKNGVFPANDATE
VLNKLHAPLRTVIRLPQTMHSGIALLEIFGGRVTDAPISKNMLSGLIADSOI
SASTOSTEALFSPSARLVYSRSGMPPRIIOAQPGEMQLQVLDGTYTKVGIIGAG
GDSITAEALAFPRKFVYSLNGKMEIIOOPRPOQPLFEGMNHVTPPIRFPDI
PAQVAVYPRMSFAGIGMRLLEYLGDMDTDSPTVETGLPVKSEETTPYRTEBAT
EGGNGCFEDDKXLQLPBSGNCNPDLEPSCMWYDHAQMLRTTMASSSPNDDT
DNATLRQSDOSRQEGQARLISPPVHLFSPVCMFQYQATGKRVLAQVREASQS
KLWLVIEDQSGWKHGRILIPSYDMEYOIVEGVIGKRGSELAIDIRISTVPLE
NCPMEIPAFGENPKVDIPIHERBEGVEIDIDEYEVDMSSSASTSGSGASPSTDEK
SMVLTLPIRLTIAMSSLGVLGATCAAGLILCTCSYSGLSRSRCTLLIENVNELYD
GLKKVMNHNOKCSSEA"

CD5
join(3518 . 3590,17760 . 17937,36403 . 36584,42688 . 42918,
43995 . 44150,46123 . 46292,48101 . 48256,60728 . 60872,
63412 . 637

overlap; the clone sequenced to the right is RP11-394E1, 200 base pair overlap. Actual start of this clone is at base position 11794 of RP11-325M10; actual end is at base position 115845 of RP11-150F11.

FEATURES

Source

Location/Qualifiers
1..116039
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"

misc_feature

/clone="RP11-150F11"
/clone_11b="RP11-11"
12630..13592
/note="CpG island (%GC=70.3, o/e=0.84, #CpGs=108)"
14041..97553
/gene="NRP2"

mRNA

join(114041..14113,28294..28471,46943..47124,53228..53458,54535..54690,56663..56832,58641..58796,71269..71413,73953..74302,76496..76640,80475..80591,83585..83725,94424..94686,96224..96320,97533..97553)
/gene="NRP2"

CDS

join(114041..14113,28294..28471,46943..47124,53228..53458,54535..54690,56663..56832,58641..58796,71269..71413,73953..74302,76496..76640,80475..80591,83585..83725,94424..94686,96224..96320,97533..97553)
/gene="NRP2"
/note="Homo sapiens neuropilin 2 (NRP2), mRNA;
H_NH0150F11.1
This gene was based on gi(4505458)
Continues as H_NH0394E01.1"

/codon_start=1
/product="unknown"
/protein_id="AA93216.1"
/db_xref="gi:62702291"

misc_feature

15626..17342
/gene="NRP2"
/note="CpG island (%GC=64.9, o/e=0.68, #CpGs=145)"

ORIGIN

Query Match 7.1%; Score 39.8; DB 8; Length 116039;
Best Local Similarity 64.8%; Pred. No. 36;

Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 4 GCCCTGCTCAGAAAAGAGAGACGACACTCAAGTACTCAACCCCATGTGACT 63
Db 77259 GCCCTCCTCGGAACAGAGAGCTGGGACACCTCGTGGCTCTACACCTCATTTGGAA 77200
QY 64 GAGAGCGAGGAGAGCTCTGAGGAGGCCCA 94
Db 77199 TGATGAGAGGTGTTCCCAAGGCTGTGGCCA 77169

RESULT 14
AC159874/c 233561 bp DNA linear HTG 01-JUL-2005
LOCUS AC159874
DEFINITION Bos taurus clone CH240-5418, *** SEQUENCING IN PROGRESS ***, 26
unordered pieces.
AC159874
AC159874

VERSION AC159874.3 GI:68267115
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Bos taurus (cow)

ORGANISM

REFERENCE

AUTHORS

Munry D, Marie, Metzker M, Lee, Abramson S, Adams C, Alder J, Allen C, Allen H, Albrooke S, Amin A, Anguiano D, Anyalebechi V, Aoyagi A, Ayodeji M, Baca E, Baden H, Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F, Biwalto K, Blair D, Blankenburg K, Blyth P, Brown M, Bryant N, Buhay C, Burch P, Burrell K, Calderon A, Cardenas V, Carter K, Cavazos I, Caesar H, Center A, Chacko J, Chavez D, Chen G, Chen R, Chen Y, Chen Z, Chu J, Cleveland C, Cockrell R, Cox C, Coyle M, Cree A, D'Souza L, Davila M, L, Davis C, Davy-Carroll L, De Anda C, Dedrich D, Delgado O, Denon S, Deramo C, Ding Y, Dinh H, Divya K, Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K, Egan A, Escotto M, Eugene C, Evans C, Falle T, Fan G, Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P, Fraser C, M, Gabisi A, Gante R, Garcia A, Garner T, Garza M, Gebregeorgis E, Geer K, Gill R, Grady M, Guerra M, Guevara W, Gunaratne P, Healand W, Hamill C, Hamilton C, Hamilton K, Harvey Y, Havlak P, Hawes A, Henderson N, Hernandez J, Hernandez R, Hines S, Hladun S, L, Hodgson A, Hognes M, Hollins B, Howell S, Huliy S, Hume J, Idelbird D, Jackson A, Jackson B, Jacob L, Jiang H, Johnson B, Johnson R, Jolivet A, Karpachy S, Kelly S, Kelly S, Khan Z, King L, Kovac C, Kwis C, Kraft C, L, Lebow H, Levan J, Lewis L, Li Z, Liu J, Liu J, Liu W, Liu Y, London P, Longacre S, Lopez J, Lorenshuwa L, Loulsegged H, Lozada R, J, Lu X, Ma J, Maheshwari M, Mahindaratne M, Mahmoud M, Malloy K, Mangum A, Mangum B, Mapa P, Martin K, Martin R, Martinez B, McWhiney S, McLeod M, P, McNeill T, Z, Meenan E, Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S, Morgan M, Morris K, Morris S, Munidasa M, Murphy M, Nair L, Nankervis C, Neal D, Newton N, Nguyen N, Norris S, Parks K, Nkoekelmen O, Okunolu G, Olarinmagaon A, Pal S, Parks K, Pasternak S, Paul H, Perez A, Perez L, Plankoch C, Plopper F, Poldexter A, Popovic D, Primus E, Pu L, L, Pizzo M, Quiroz J, Rachlin E, Reeves K, Regier M, A, Reigh R, Reilly B, Reilly M, Ren Y, Reuter M, Richard S, Riggs F, Rivers C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, J, Sanders W, Saverly G, Scherer S, Scott G, Shasman S, Shen H, Shetty J, Shvartsbeyn A, Sisson I, Sitter C, D, Smajs D, Sneed A, Sodergren E, Song X, Z, Sorelle R, Soosa J, Steidle M, Strong R, Sutton A, Svatek A, Tabor P, Taylor C, Taylor T, Thomas N, Thomas S, Tingey A, Trejos Z, Umani K, Valae R, Vera V, Villaseana D, Waldron J, Walker B, Wang J, Wang O, Wang S, Warren J, Warren R, Wei K, White F, Williams G, Willson R, Wlezyk R, Wooden H, Worley K, Wright D, Wright R, Wu J, Yakub S, Yen J, Yoon L, Yoon V, Yu P, Zhang J, Zhou J, Zhou X, Zhou S, Dunn D, von Niederhausen A, Weis R, Smith D, R, Holt R, A, Smith H, O, Weinstock G, and Gibbs R, A.

Direct Submission
Unpublished
2 (bases 1 to 233561)
REFERENCE
Worley K, C.
Direct Submission
Submitted (13-APR-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 233561)
REFERENCE
Cow Genome Sequencing Consortium.
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 28, 2005 this sequence version replaced gi:62543289.
The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FDDIS

Center clone name: CH240-5418

----- Summary Statistics

Assembly program: Atlas 3.0/

Consensus quality: 226744 bases at least Q40

Consensus quality: 228422 bases at least Q30

Consensus quality: 229869 bases at least Q20

Estimated insert size: 229962; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/doc/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 26 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

```

* 1 4475: contig of 4475 bp in length
* 4476 4794: gap of 319 bp
* 4795 10004: contig of 5210 bp in length
* 10005 10054: gap of 50 bp
* 10055 14989: contig of 4935 bp in length
* 14990 15039: gap of 50 bp
* 15040 23192: contig of 8153 bp in length
* 23193 23242: gap of 50 bp
* 23243 27656: contig of 4414 bp in length
* 27657 27706: gap of 50 bp
* 27707 32345: contig of 4639 bp in length
* 32346 32395: gap of 50 bp
* 32396 80065: contig of 47670 bp in length
* 80066 80258: gap of 193 bp
* 80259 95526: contig of 15268 bp in length
* 95527 95576: gap of 50 bp
* 95577 11151: contig of 15575 bp in length
* 11152 111201: gap of 50 bp
* 111202 127875: contig of 16674 bp in length
* 127876 127925: gap of 50 bp
* 127926 130975: contig of 3050 bp in length
* 130976 131064: gap of 89 bp
* 131065 136494: contig of 5430 bp in length
* 136495 136594: gap of unknown length
* 136595 140341: contig of 3747 bp in length
* 140342 140391: gap of 50 bp
* 140392 141698: contig of 1307 bp in length
* 141699 141748: gap of 50 bp
* 141749 151343: contig of 9595 bp in length
* 151344 151393: gap of 50 bp
* 151394 154790: contig of 3397 bp in length
* 154791 154840: gap of 50 bp
* 154841 166899: contig of 12059 bp in length
* 166900 166949: gap of 50 bp
* 166950 172623: contig of 5674 bp in length
* 172624 172673: gap of 50 bp

```

```

* 172674 185758: contig of 13085 bp in length
* 185759 185808: gap of 50 bp
* 185809 210143: contig of 24335 bp in length
* 210144 210193: gap of 50 bp
* 210194 226003: contig of 15810 bp in length
* 226004 226103: gap of unknown length
* 226104 227673: contig of 1570 bp in length
* 227674 227774: gap of unknown length
* 227775 229134: contig of 1361 bp in length
* 229135 229234: gap of unknown length
* 229235 230328: contig of 1093 bp in length
* 230329 230427: gap of unknown length
* 230428 231641: contig of 1214 bp in length
* 231642 231741: gap of unknown length
* 231742 233561: contig of 1820 bp in length.

```

FEATURES

source

```

1. .233561
   /organism="Bos taurus"
   /mol_type="genomic DNA"
   /db_xref="taxon:9913"
   /clone="CH240-5418"
   4476..4794
   /estimated_length=319
   10005..10054
   /estimated_length=50
   14990..15039
   /estimated_length=50
   23193..23242
   /estimated_length=50
   27657..27706
   /estimated_length=50
   32346..32395
   /estimated_length=50
   80066..80258
   /estimated_length=193
   95527..95576
   /estimated_length=50
   11152..111201
   /estimated_length=50
   127876..127925
   /estimated_length=50
   130976..131064
   /estimated_length=89
   136495..136594
   /estimated_length=unknown
   140342..140391
   /estimated_length=50
   141699..141748
   /estimated_length=50
   151344..151393

```

Query Match 7.1%; Score 39.8; DB 14; Length 233561;

Best Local Similarity 60.7%; Pred. No. 33; Mismatches 42; Indels 0; Gaps 0;

Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

QY 274 GCGTTGACAGGATTGGGGCTTCACATCACAGGATCCATCCAGTACCAAGCCCAAGTGGC 333
      |||||
DB 125658 GCGTCCAGGAGATTAGCTGCTGTGAGAGCTGGCCCCACAGTCCACGACATGCCCGC 125599

```

```

QY 334 TGCAGAGGTCCTTCGACGTCATGAACCAAGGAGCTTGGGAAC 380
      |||||
DB 125598 TCTGGCAGATCTCTCAGGAGTCAGTCTGCCAGAGTGGGGGTGGGAGC 125552

```

RESULT 15

AC124345

LOCUS

DEFINITION Mus musculus BAC clone RP24-459N19 from chromosome 15, complete

sequence.

AC124345

VERSION

AC124345.4

KEYWORDS

SOURCE

184532 bp DNA linear ROD 27-NOV-2003
Mus musculus BAC clone RP24-459N19 from chromosome 15, complete
GI:37361090
Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 184532)
TITLE Dauphin, S. and Hakenson, W.
JOURNAL The sequence of Mus musculus BAC clone RP24-459N19
REFERENCE
AUTHORS 2 (bases 1 to 184532)
TITLE Wilson, R.
JOURNAL Sequencing of Mus musculus
REFERENCE
AUTHORS 3 (bases 1 to 184532)
TITLE McPherson, J.D. and Waterston, R.H.
JOURNAL Direct Submision
REFERENCE
AUTHORS 4 (bases 1 to 184532)
TITLE McPherson, J.D. and Waterston, R.H.
JOURNAL Submitted (14-JUN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS 5 (bases 1 to 184532)
TITLE McPherson, J.D. and Waterston, R.H.
JOURNAL Submitted (19-JAN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS 6 (bases 1 to 184532)
TITLE Wilson, R.K.
JOURNAL Direct Submision
REFERENCE
AUTHORS Submitted (02-OCT-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Oct 2, 2003 this sequence version replaced gi:27803375.

Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu
Contact: submissions@waterston.wustl.edu

Summary Statistics
Center project name: M_BB0459N19

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-24 BAC library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC021667.
Location/Qualifiers
1. .184532
/organism="Mus musculus"

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="15"
/map="15"
/clone="RP24-459N19"
/clone_id="RPCI-24"
2657. .2749
/rpt_family="L1"
repeat_region
9553. .9657
/rpt_family="Alu"
repeat_region
13735. .13788
/rpt_family="ERV1"
repeat_region
16979. .17050
/rpt_family="B4"
misc_feature
19402. .20528
/note="CpG island (GC=61.5, o/e=0.81, #CpGs=82)"
27391. .27618
/note="Sequence derived from one plasmid subclone."
35924. .36046
/rpt_family="ERVK"
repeat_region
36205. .36316
/note="Sequence derived from PCR product of project DNA."
44255. .44412
/rpt_family="ERVK"
misc_feature
67067. .68227
/note="CpG island (GC=61.0, o/e=0.72, #CpGs=74)"
79424. .80411
/note="CpG island (GC=63.3, o/e=0.94, #CpGs=81)"
84246. .84488
/rpt_family="ERVK"
repeat_region
85290. .85359
/rpt_family="ERV1"
repeat_region
103839. .103921
/rpt_family="ERV1"
repeat_region
107737. .107995
/rpt_family="MER1_type"
112536. .113536
/note="CpG island (GC=71.3, o/e=0.75, #CpGs=108)"
116901. .116956
/rpt_family="ERV1"
repeat_region
122809. .122934
/rpt_family="MER2_type"
122940. .123121
/rpt_family="L1"
repeat_region
125910. .125991
/rpt_family="MALR"
repeat_region
126174. .126343
/rpt_family="B4"
repeat_region
126376. .126499
/rpt_family="Alu"
repeat_region
126377. .126524
/rpt_family="B4"
repeat_region
127557. .127594
/rpt_family="B2"
repeat_region
128080. .128211
/rpt_family="Alu"
repeat_region
128101. .128291
/rpt_family="B4"
repeat_region
128619. .128690
/rpt_family="B2"
repeat_region
130435. .130504
/rpt_family="RD"
repeat_region
130519. .130771
/rpt_family="Alu"
repeat_region
130530. .130559
/rpt_family="B4"
repeat_region
130560. .130663
/rpt_family="Alu"
repeat_region
130664. .130810
/rpt_family="B4"
repeat_region
130744. .130804
/rpt_family="MALR"
repeat_region
130805. .130890

```
repeat_region /rpt_family="Alu"
130831. .130908 /rpt_family="MALR"
repeat_region 131637. .131710
/rpt_family="I1"
repeat_region 132036. .132173
/rpt_family="MIR"
repeat_region 132229. .132420
/rpt_family="I1"
repeat_region 132589. .132719
/rpt_family="Alu"
repeat_region 132625. .132739
/rpt_family="MALR"
repeat_region 132982. .133110
/rpt_family="Tc2"
repeat_region 133154. .133332
/rpt_family="Alu"
repeat_region 133399. .133593
/rpt_family="B2"
repeat_region 133975. .134204
/rpt_family="B4"
repeat_region 134214. .134325
/rpt_family="Alu"
repeat_region 134638. .134710
/rpt_family="cRNA-Ala-GCA"
repeat_region 134740. .134925
/rpt_family="I1"
repeat_region 135953. .136036
/rpt_family="B4"
repeat_region 136118. .136259
/rpt_family="Alu"
repeat_region 136392. .136569
/rpt_family="I1"
repeat_region 136705. .136859
/rpt_family="ERV1"
repeat_region 137058. .137195
/rpt_family="Alu"
repeat_region 137500. .137647
/rpt_family="Alu"
repeat_region 137668. .137764
/rpt_family="Alu"
repeat_region 138021. .138147
/rpt_family="MIR"
repeat_region 138278. .138348
/rpt_family="ID"
repeat_region 139137. .139318
/rpt_family="B2"
```

```
Query Match 6.9%; Score 38.8; DB 9; Length 184532;
Best Local Similarity 51.8%; Pred. No. 59;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 364 AGGGAGGCTTGGGAAACCATCTGAAGGCGATGCTTGATTGAGAGGCTGGGCG 423
DB 81285 ATGTATGCTAAGGGAACATCTTGAAGAGCTTTTTTTTTTATGATTGGGAGCT 81344
QY 424 TGGGCTGGGCAAGGCACACAGGCTGAGTCAGAGCCAGAGCAGGAAGCTGCTCCCAAC 483
DB 81345 TAGGTGGGTAGAGGTGTCTATCTTAATCCAAATCAAAAGTAGCAGGCTAGGGCTGA 81404
QY 484 ACTGCCCGCGCGCTCTGCATCGAGTCTCTCGCCACCTGAGAACAGCC 533
DB 81405 AGTCCCAAGCCTTTCACTCTATTTCTTCTCTACATCTCATGTCACACC 81454
```

Search completed: March 25, 2006, 17:51:42
Job time : 3016.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 15:57:57 ; Search time 472 Seconds
(without alignment)
7963.743 Million cell updates/sec

Title: SEQ1-69C

Perfect score: 564

Sequence: 1 gcagcccttcctcagaaac.....agtgcgcttcctcgacttc 564

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: geneeqn1980s:*
- 2: geneeqn1980s:*
- 3: geneeqn2000s:*
- 4: geneeqn2001as:*
- 5: geneeqn2001bs:*
- 6: geneeqn2002as:*
- 7: geneeqn2002bs:*
- 8: geneeqn2003as:*
- 9: geneeqn2003bs:*
- 10: geneeqn2003cs:*
- 11: geneeqn2003ds:*
- 12: geneeqn2004as:*
- 13: geneeqn2004bs:*
- 14: geneeqn2005as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	563	99.8	564	13 ADT97901	Adt97901 Human fib
2	553	98.0	29340	13 ABD33482	Abd33482 Human can
3	553	98.0	57501	14 ADZ13482	Adz13482 Human can
4	38.2	6.8	2161	8 ABX34807	Abx34807 Human mdd
5	38.2	6.8	2777	11 ADM01893	Adm01893 Human cdn
6	37.2	6.6	712	6 ABK34393	Abk34393 Human cdn
7	37.2	6.6	1957	4 ABK42424	Abk42424 Genomic s
8	37.2	6.6	1957	4 AAL07453	Aal07453 Human rep
9	37.2	6.6	1957	4 ADB60580	Adb60580 Connectiv
10	37	6.6	1785	6 AAS19511	Aas19511 Reference
11	36.6	6.5	550	9 ACH25075	Ach25075 Human adu
12	36.4	6.5	76180	13 ABD33385	Abd33385 Human can
13	36.2	6.4	138203	14 ADZ51725	Adz51725 FR-008 po
14	36	6.4	546	13 ADD56739	Adg56739 Novel can
15	35.8	6.3	704	2 AAO94044	Aag94044 Human ALD
16	35.8	6.3	1263	6 ABO90455	Abg90455 M. capsul
17	35.6	6.3	568	5 ABV51994	Abv51994 Human pro
18	35.2	6.2	954	14 ACT70503	Act70503 M. xanthu
19	35.2	6.2	4050	14 ACT64141	Act64141 M. xanthu

20	35	6.2	1338	6 ABL58452	Ab158452 Human pel
21	35	6.2	1338	12 ADP48674	Adp48674 Human pel
22	35	6.2	1797	4 ABK42425	Abk42425 Genomic s
23	35	6.2	1797	3 AAL07455	Aal07455 Human rep
24	35	6.2	1797	9 ADB60581	Adb60581 Connectiv
25	35	6.2	2000	8 ADA71938	Ada71938 Rice gene
26	34.8	6.2	792	13 ADR62531	Adr62531 Cotton cd
27	34.6	6.1	40742	4 AAK68089	Aak68089 Human imm
28	34.6	6.1	215980	6 AAK79886	Aak79886 Human imm
29	34.6	6.1	215980	6 AAL38337	Aal38337 Complemen
30	34.6	6.1	220895	6 AAK84798	Aak84798 Human cdn
31	34.6	6.1	220895	13 ADR52737	Adr52737 Drug ther
32	34.4	6.1	1988	13 ADS56480	Ads56480 Bacterial
33	34.4	6.1	2000	8 ADA71938	Ada71938 Rice gene
34	34.4	6.1	5051	13 ADQ39151	Adq39151 Human SNP
35	34.4	6.1	5232	13 ADQ39152	Adq39152 Human SNP
36	34.4	6.1	92638	6 ABO88096	Abg88096 Human oet
37	34.2	6.1	756	2 AAX85016	Aax85016 Human sec
38	34.2	6.1	756	6 ACD18942	Act18942 Novel hum
39	34.2	6.1	756	12 ADG78333	Adg78333 Human sec
40	34.2	6.1	756	12 ADN60624	Adn60624 Human sec
41	34.2	6.1	349980	6 ABO81844	Abg81844 Bifidobac
42	34	6.0	536	10 ADB68842	Adb68842 Minorily
43	34	6.0	593	12 ACH71544	Ach71544 Human gen
44	34	6.0	1028	10 ADD84538	Add84538 121P1 v
45	34	6.0	1028	10 ADD84540	Add84540 121P1 v

ALIGNMENTS

RESULT 1	ADT97901	ID	ADT97901 standard; DNA; 564 BP.
XX	ADT97901;	DT	27-JAN-2005 (first entry)
XX	DE	Human fibroblast growth factor (FGF)-3 promoter 5' proximal region.	
XX	KX	Human; ds; fibroblast growth factor-3; FGF-3; promoter; SNP;	
XX	KW	single nucleotide polymorphism; cancer; oesophageal cancer; 5'UTR;	
XX	KW	breast cancer; ovarian cancer; prostate cancer; head and neck cancer;	
XX	KW	oesophageal squamous cell carcinoma.	
OS	OS	Homo sapiens.	
XX	XX	Key	
XX	XX	Location/Qualifiers	
XX	XX	replace(69,'Y')	
XX	XX	/*tag= a	
XX	XX	/standard name= "Single nucleotide polymorphism"	
XX	XX	/note= "This SNP is specifically claimed in claim 6"	
XX	XX	US2004219582-A1.	
XX	XX	04-NOV-2004.	
XX	XX	11-MAR-2004; 2004US-00798652.	
XX	XX	17-MAR-2003; 2003US-0455689P.	
XX	XX	(GUOY/) GUO Y.	
XX	XX	Guo Y;	
XX	XX	WPI; 2004-794435/78.	
XX	XX	Novel isolated nucleic acid molecule having single nucleotide	
XX	XX	polymorphism in upstream untranslated region of fibroblast growth factor	
XX	XX	3 gene; useful for assessing related susceptibility of mammal to cancer.	
XX	XX	Claim 1; SEQ ID NO 1; 15pp; English.	


```
Db 3746 CTGGTGCCAGCCTCTGGAAGTCTGGAGTGAAGCATGTTTCCATTAAAGAAAGTGTG 3805
Qy 241 GCGGCGCATGCCCCCAAGCTTGGACACTCACTGCTTTGGAGGGTGGGCTTCCAGTC 300
Db 3806 GCGGCGCATGCCCCCAAGCTTGGACACTCACTGCTTTGGAGGGTGGGCTTCCAGTC 3865
Qy 301 ACAGGGTCCATCCACGTACAGCCAGGAGTGGCTGACAGAGGTCCTCCGACGTATGAAA 360
Db 3866 ACAGGGTCCATCCACGTACAGCCAGGAGTGGCTGACAGAGGTCCTCCGACGTATGAAA 3925
Qy 361 CCAAGGGAGGCTTGGGAAACCATCTGAAGGGCATGGCTTTGATTAGTAGAGGGTGG 420
Db 3926 CCAAGGGAGGCTTGGGAAACCATCTGAAGGGCATGGCTTTGATTAGTAGAGGGTGG 3985
Qy 421 GGCTGGGCTGGGCAAGGCGACAGGCTGAGTCAAGAGCCAGAGGCAAGCTGCTCC 480
Db 3986 GGCTGGGCTGGGCAAGGCGACAGGCTGAGTCAAGAGCCAGAGGCAAGCTGCTCC 4045
Qy 481 A-GCACTGCGCGCGCTCTGCGATGACATGCTCTCTGGCCACTGAGAAAGCCTGTAGA 539
Db 4046 AGGCACTGCGCGCGCTCTGCGATGACATGCTCTCTGGCCACTGAGAAAGCCTGTAGA 4105
Qy 540 GAGGCACTGCGCGCTCTTTCGAACTTC 564
Db 4106 GAGGCACTGCGCGCTCTTTCGAACTTC 4130
```

```
RESULT 3
AD213482
ID AD213482 standard; DNA; 57501 BP.
XX
AC AD213482;
XX
DT 16-JUN-2005 (first entry)
XX
DE Human cancer-associated genomic DNA #85.
XX
KM Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
XX cytostatic; gene; db.
XX
OS Homo sapiens.
XX
PN W0205031001-A2.
XX
PD 07-APR-2005.
XX
PF 23-SEP-2004; 2004WO-US031617.
XX
PR 23-SEP-2003; 2003US-00669920.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2005-273395/28.
XX
PT Nucleic acid array useful for detecting cancer associated nucleic acid,
XX comprises two or more nucleic acid probes.
XX
PS Disclosure; SEQ ID NO 1002; 198pp; English.
XX
XX The invention relates to a nucleic acid array for detecting a cancer
XX associated (CA) nucleic acid, comprising two or more nucleic acid probes.
XX The invention also relates to a peptide array comprising two or more
XX isolated polypeptides encoded by a CA nucleic acid sequence, a compound
XX that binds to a polypeptide, an isolated antibody or its fragment which
XX binds to a polypeptide, which is prepared by immunizing a host animal
XX with a composition comprising the polypeptide or its antigen binding
XX fragment and collecting cells from the host expressing antibodies against
XX the antigen or its antigen binding fragment, a composition comprising the
XX antibody and a carrier, a method of screening for anticancer activity, a
XX method of detecting a CA nucleic acid, a method of diagnosing cancer, a
```

```
CC method of treating cancer and a method of inhibiting expression of a CA
CC nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
CC nucleic acids. The antibody is useful for detecting the presence or
CC absence of cancer cells in an individual which involves contacting cells
CC from the individual with the antibody and detecting a complex of a CA
CC protein from the cancer cells and the antibody, where the detection of
CC the complex correlates with the presence of cancer cells in the
CC individual. The composition is useful for inhibiting growth of cancer
CC cells in an individual or for delivering a therapeutic agent to cancer
CC cells in an individual. The invention is also useful for diagnosing
CC cancer, for treating cancer and for inhibiting expression of a CA gene in
CC a cell. This sequence represents human cancer-associated genomic DNA of
CC the invention.
XX
```

Sequence 57501 BP; 13274 A; 15989 C; 15611 G; 12627 T; 0 U; 0 Other;

Query Match 98.0%; Score 553; DB 14; Length 57501;
Best Local Similarity 99.8%; Pred. No. 1,7e-138;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy 1 GCAAGCCCTGCTCAGAAACAGAAAGGACAGCACTCAAGGACTCAAGGCTGAGTGC 60
Db 3716 GCAAGCCCTGCTCAGAAACAGAAAGGACAGCACTCAAGGACTCAAGGCTGAGTGC 3775
Qy 61 GCTGAGGCGAGGAGGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 3776 GCTGAGGCGAGGAGGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3835
Qy 121 GGGGCTTGGCCATGAGTGGGCAAGGAGGCTGATACAGCTTACTGAGTACATGAGTCC 180
Db 3836 GGGGCTTGGCCATGAGTGGGCAAGGAGGCTGATACAGCTTACTGAGTACATGAGTCC 3895
Qy 181 CTGGTGCAGGCTCTGGAATCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 240
Db 3896 CTGGTGCAGGCTCTGGAATCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3955
Qy 241 GCGGAGCATGCCCCCAAGCTTGCACACATGACATGCTCTTTCAGAGGAGTGGGCTTCCAGTC 300
Db 3956 GCGGAGCATGCCCCCAAGCTTGCACACATGACATGCTCTTTCAGAGGAGTGGGCTTCCAGTC 4015
Qy 301 ACAAGGTCCTCATCAGCTACAGCCAGGAGTGGCTGCAAGAGTCCCTGCGATCATGAAA 360
Db 4016 ACAAGGTCCTCATCAGCTACAGCCAGGAGTGGCTGCAAGAGTCCCTGCGATCATGAAA 4075
Qy 361 CCAAGGAGGCTTGGGAAACCATCATCTGAAGGGCATGCTTTGATTAGTAGAAGGGTGG 420
Db 4076 CCAAGGAGGCTTGGGAAACCATCATCTGAAGGGCATGCTTTGATTAGTAGAAGGGTGG 4135
Qy 421 GGCTGGGCTGGGCAAGGCGACAGGCTGAGTCAAGAGCCAGAGGCAAGAGGCTGATCCCC 480
Db 4136 GGCTGGGCTGGGCAAGGCGACAGGCTGAGTCAAGAGCCAGAGGCAAGAGGCTGATCCCC 4195
Qy 481 A-GCACTGCGCGCGCTCTGCGATGACATGCTCTTGGCCACTGAGAAAGCCTGTAGA 539
Db 4196 AGGCACTGCGCGCGCTCTGCGATGACATGCTCTTGGCCACTGAGAAAGCCTGTAGA 4255
Qy 540 GAGGCACTGCGCGCTTTCGAACTTC 564
Db 4256 GAGGCACTGCGCGCTTTCGAACTTC 4280
```

```
RESULT 4
ABX34807/c
ID ABX34807 standard; cDNA; 2161 BP.
XX
AC ABX34807;
XX
DT 13-FEB-2003 (first entry)
XX
DE Human mdlt cDNA SEQ ID 368.
XX
XX MDDT; human; disease detection and treatment molecule polypeptide;
XX anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
```


QY	129	ACTCTGGAAGTAGCAATGTTTCCCATTTAAGAAAGTGTGGCCGGCCATGCCCCCA	258
DB	225	ATTGAGCAGAGGACCACTTGTTCCTTGTATGGACCCAGGACCCACGAAGCCGATGGC	166
OY	259	CGTTGCACACTCACTGCTCTTTCGAGGGTTGG	289
DB	165	CATGACAAAGGCGCCGGTATGATGACGAG	135
RESULT 6			
ID	ABK34393/C		
XX	ABK34393 standard; cDNA; 712 BP.		
XX	ABK34393;		
AC			
XX	08-MAY-2002 (first entry)		
DT			
DE			
XX	Human cDNA for novel secreted protein, SEQ ID 162.		
KM	Human; 68; gene; secreted protein; immune deficiency; viral infection;		
KM	bacterial infection; fungal infection; autoimmune disorder; burn;		
KM	rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;		
KM	diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;		
KM	Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;		
KM	coagulation disorder; haemophilia; inflammatory disorder; ulcer;		
KM	tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;		
KM	lymphoid cell deficiency.		
OS			
XX	Homo sapiens.		
PN	W0200177230-A2.		
XX	18-OCT-2001.		
PD			
XX	29-MAR-2001; 2001WO-US010295.		
PF			
PR	06-APR-2000; 2000US-0194941P.		
XX			
PA	(GENY) GENETICS INST INC.		
PI	Mong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ,		
PI	Gulukota K, Graham JR;		
DR	WPI; 2002-179323/23.		
PT			
PT	Six hundred and twenty five polynucleotides derived from a variety of		
PT	human tissue sources which encode secreted proteins, useful for treating		
XX	immune deficiencies and disorders such as autoimmune disorders.		
XX			
XX	Claim 1; Page 114; 339pp; English.		
PS			
XX	The invention relates to 625 polynucleotides which have been derived from		
CC	a variety of human tissue sources and which encode novel secreted		
CC	proteins, their complements and sequences that hybridise to them. Also		
CC	included are a vector comprising the polynucleotide, a host cell		
CC	transformed with the vector, the proteins encoded by the polynucleotides,		
CC	antibodies that bind to the proteins and identification of modulators of		
CC	the proteins or the expression of the polynucleotide. The polynucleotides		
CC	can be used as probes for the identification and isolation of full length		
CC	cDNA and genomic DNA. The polynucleotides and proteins can also be used		
CC	as nutritional supplements. The protein is useful in the treatment of		
CC	various immune deficiencies and disorders such as viral infections,		
CC	bacterial infections, fungal infections, autoimmune disorders (e.g.		
CC	rheumatoid arthritis, multiple sclerosis, autoimmune thyroiditis and		
CC	diabetes) and allergic reactions and conditions (e.g. asthma). They are		
CC	also useful for treating neurodegenerative diseases (e.g. Alzheimer's		
CC	disease, Parkinson's disease), liver fibrosis, coagulation disorders		
CC	(e.g. haemophilia), inflammatory disorders (e.g. Crohn's disease) and		
CC	tumours. They are also useful for tissue regeneration, for wound healing		
CC	and in the treatment of burns, incisions and ulcers. The proteins are		
CC	also useful for regulating haematopoiesis, for treating myeloid or		
CC	lymphoid cell deficiencies. The present sequence is one of the 625 cDNA		
CC	sequences encoding a secreted protein		

Seq	Sequence	712 BP, 141 A, 159 C, 205 G, 207 T, 0 U, 0 Other;
QY	Query Match	6.6%; Score 37.2; DB 6; Length 712;
Db	Best Local Similarity	56.6%; Pred. No. 7.1;
QY	Matches 69; Conservative	0; Mismatches 53; Indels 0; Gaps 0
Db	500	AGGGAATCCACCAACAGAGAGTGTGACAGCGGCTGTGGCAATGCCAGGACA 441
QY	371 CT 372	
Db	440 CT 439	
RESULT 7		
ID	ABK42424	
AC	ABK42424 standard; DNA; 1957 BP.	
DT	21-MAY-2002 (first entry)	
DE	Genomic sequence #333 encoding novel human connective tissue polypeptide.	
XX	Human; connective tissue related disorder; cancer; gene therapy;	
XX	cytostatic; gene; de.	
XX	Homo sapiens.	
XX	WO200155343-A1.	
XX	02-AUG-2001.	
XX	17-JAN-2001; 2001WO-US001322.	
PR	31-JAN-2000; 2000US-0179065P.	
PR	04-FEB-2000; 2000US-0180628P.	
PR	24-FEB-2000; 2000US-0184664P.	
PR	02-MAR-2000; 2000US-0186350P.	
PR	16-MAR-2000; 2000US-0189874P.	
PR	17-MAR-2000; 2000US-0190076P.	
PR	18-APR-2000; 2000US-0198123P.	
PR	19-MAY-2000; 2000US-0205515P.	
PR	07-JUN-2000; 2000US-0209467P.	
PR	28-JUN-2000; 2000US-0214886P.	
PR	30-JUN-2000; 2000US-0215135P.	
PR	07-JUL-2000; 2000US-0216647P.	
PR	07-JUL-2000; 2000US-0216880P.	
PR	11-JUL-2000; 2000US-0217487P.	
PR	11-JUL-2000; 2000US-0217496P.	
PR	14-JUL-2000; 2000US-0218290P.	
PR	26-JUL-2000; 2000US-0220963P.	
PR	26-JUL-2000; 2000US-0220964P.	
PR	14-AUG-2000; 2000US-0224518P.	
PR	14-AUG-2000; 2000US-0224519P.	
PR	14-AUG-2000; 2000US-0225213P.	
PR	14-AUG-2000; 2000US-0225214P.	
PR	14-AUG-2000; 2000US-0225266P.	
PR	14-AUG-2000; 2000US-0225267P.	
PR	14-AUG-2000; 2000US-0225268P.	
PR	14-AUG-2000; 2000US-0225447P.	
PR	14-AUG-2000; 2000US-0225757P.	
PR	14-AUG-2000; 2000US-0225758P.	
PR	14-AUG-2000; 2000US-0225759P.	
PR	18-AUG-2000; 2000US-0226279P.	
PR	22-AUG-2000; 2000US-0226281P.	

RESULT 8
AAL07453
ID AAL07453 standard; DNA; 1957 BP.
XX AAL07453;
AC
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 10141.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246517P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

[illegible]

PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
PR	17-JAN-2001;	2001US-00764847.
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Ruben SM, Barnash SC;	
DR	WPI; 2003-634869/60.	
XX	P-PSDB; ADB59454.	
PT	New connective tissue-related polypeptides and polynucleotides, useful	
PT	for treating, preventing and/or prognosing e.g. disorders of connective	
PT	tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or	
PT	neoplasias.	
PS	Disclosure; SEQ ID NO 1311; 248bp; English.	
XX		
XX	The invention describes an isolated nucleic acid molecule (1), which	
CC	comprises a sequence that is at least 95 % identical to a connective	
CC	tissue-related polynucleotide encoding connective tissue antigens (CTA).	
CC	The polypeptide or polynucleotide is useful for preventing, treating, or	
CC	ameliorating medical conditions in a mammal. The connective tissue	
CC	polypeptides, polynucleotides and antibodies are particularly useful for	
CC	treating, preventing and/or prognosing disorders of connective tissues	
CC	(e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,	
CC	scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or	
CC	neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.	
CC	Alzheimer's disease, or Parkinson's disease), cardiovascular diseases	
CC	(e.g. atherosclerosis, myocarditis or cardiopulmonary bypass	
CC	complications), autoimmune diseases (e.g. systemic lupus erythematosus,	
CC	rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.	
Query Match	6.6%; Score 37.2; DB 9; Length 1957;	
Best Local Similarity	56.6%; Pred. No. 9.3;	
Matches	69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;	
Oy	251 CCCCCCAAGTTCGACACTGCTGCGCTTGGACGGGCTTCAGTCACAGGGTCC	310
Db	165 CCGCCCAAGCTGGCTCACTTCTTGCCCTGGAGGAAAGGGCCGAGGGGCTTCTGTTCC	224
Oy	311 ATCCACGTACCAAGCCCAAGTGGCTGTCAGAAAGTCCCTGCACTGATGAAACCAAGGAGG	370
Db	225 AGGGAATCCACACACAGGAGTGTGACAGAGCGGTGGCTGGCAGTCAGTCCAGGAACA	284
Oy	371 CT 372	
Db	285 CT 286	
RESULT 10		
AA519511		
DB ID	AA519511 standard; DNA; 1785 BP.	
XX	AA519511;	
AC		
XX	26-MAR-2002 (first entry)	
DT		
XX	Reference sequence for human MPL gene exons 9-10.	
DS		
XX	Human; single nucleotide polymorphism; SNP; MPL; chromosome 1p34;	
KW	myeloproliferative leukaemia virus oncogene; haplotyping; genotyping;	
KM	congenital amegakaryocytic thrombocytopaenia; CMT; ds.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	exon	614..773
FT		/*Cag= a
FT		/number= 9
FT	Intron	774..1027
FT		/*Cag= b

```
PT /number= 9
FT exon 1028. .1124
FT /*tag= c
FT /number= 10
XX
XX WO200179232-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US012301.
XX
XX 14-APR-2000; 2000US-0197839P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Chew A, Choi YJ, Koshy B, Stephens JC;
XX
XX WPI; 2002-055251/07.
XX
XX P-PSDB; AAU09495.
XX
XX Nucleotide polymorphisms in the human myeloproliferative leukemia virus
XX oncogene (MPL) gene, useful for studying the function of and expressing
XX MPL protein for use in screening drugs for treating diseases related to
XX MPL activity.
XX
XX Claim 27; Fig 3; 85pp; English.
XX
XX The present invention relates to novel single nucleotide polymorphisms
XX (SNPs) in the human myeloproliferative leukemia virus oncogene (MPL)
XX gene located on chromosome 1p34, and methods for haplotyping and/or
XX genotyping the MPL gene. The methods of the invention make use of allele-
XX specific oligonucleotides (ASOs) as probes and primers and/or primer-
XX extension oligonucleotides for detecting MPL gene polymorphisms. The
XX polymorphisms and screened compounds are useful for the treatment of
XX diseases associated with MPL activity, such as congenital amegakaryocytic
XX thrombocytopaenia (CAMT). The present sequence represents a reference
XX sequence for human MPL gene exons 9-10. Note: This sequence encodes for
XX only part of the MPL protein shown in AAU09495
XX
XX Sequence 1785 BP; 317 A; 562 C; 500 G; 403 T; 0 U; 3 Other;
XX
XX Query Match 6.6%; Score 37; DB 6; Length 1785;
XX Best Local Similarity 62.4%; Pred. No. 10;
XX Matches 58; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
XX
XX QY 53 CCCATGCTGCTGAGGCGAGGAGCCTCTGAGGAGGCGGCGAGCCCTCAAGTGG 112
XX DB 913 CGCAGGTGCGCCGACGTCCCGAGGGGCGGCGAGGGCGAGTGAAGGCTGAGTGA 972
XX
XX QY 113 TGACGCGAGGGGTCTTGCCATGTTGGGCAAGG 145
XX DB 973 TGAAGGCGGGGCTCCGGCCCGGTGGGCGAAG 1005
XX
XX RESULT 11
XX ACH25075/c
XX ID ACH25075 standard; cDNA; 550 BP.
XX
XX ACH25075;
XX
XX 13-OCT-2003 (first entry)
XX
XX Human adult ovary cDNA #3455.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
```

```
PF 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DEMA/) DEMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Dmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 12287; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations,
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antibodies specific for it. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?docID=20030073623
XX
XX Sequence 550 BP; 112 A; 142 C; 159 G; 132 T; 0 U; 5 Other;
XX
XX Query Match 6.5%; Score 36.6; DB 9; Length 550;
XX Best Local Similarity 52.1%; Pred. No. 9.6;
XX Matches 75; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
XX
XX QY 30 CAGCACCTACAGGTGACTACCCCATGTGCTGAGGAGCGAGGAGCCTCTGAGGAG 89
XX DB 206 CAGCAGGCAATTGCGGCTCCATTATTCAGGCTGAAGGCGAAGTACCCCGCAGGCTC 147
XX
XX QY 90 GGCAGGGCAGCCGCTCAAGTGGTGACGCGCAGGGGTCTTGCCATGTTGGGCAAGGGGCT 149
XX DB 146 GCGCTCGCATGAGATGGGAGCGATTAAGAAAGGNNTTCCNAAGGTTGTACAAGTGTCT 87
XX
XX QY 150 GCATACAGCTTACTCAAGTGAAT 173
XX DB 86 TCGAACCTTTNTTTTGTGTAAT 63
XX
XX RESULT 12
XX ABD33385/c
XX ID ABD33385 standard; DNA; 76180 BP.
XX
XX ABD33385;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human cancer-associated (CA) gene HD07-071.
XX
XX Human cancer-associated protein; CAP; cancer-associated gene; CA; gene;
XX ds; cancer; cytostatic.
XX
XX Homo sapiens.
XX
```

PN	M02004058146-A2.
PD	15-JUL-2004 .
XX	
XX	
PF	15-DEC-2003; 2003WO-US040081.
XX	
PR	17-DEC-2002; 2002US-00322281.
XX	
PA	(SAGR-) SAGRES DISCOVERY INC.
XX	
PI	Morris DW, Malandro MS;
DR	WPI; 2004-499109/47.
XX	
PT	Novel human cancer associated protein encoded within open reading frame
PT	of cancer associated gene, useful as targets for diagnosing cancer.
XX	
PS	Claim 16; SEQ ID NO 492; 182bp; English.
CC	The invention relates to cancer-associated proteins (CAP) and the cancer-
CC	associated (CA) nucleic acids encoding them. The invention also relates
CC	to a method for treating cancers involving administering to a patient an
CC	inhibitor of CAP, and a method of screening for anticancer activity in a
CC	potential drug involving providing a cell that expresses a CA gene,
CC	contacting a tissue sample derived from a cancer cell with an anticancer
CC	drug candidate and monitoring the effect of the anticancer drug candidate
CC	on expression of the CA gene. The CAP proteins are useful for detecting
CC	cancer associated with expression of a CAP protein in a test cell sample
CC	and for screening for a bioactive agent capable of modulating the
CC	activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC	cancer, involving determining the expression of a CA nucleic acid in a
CC	tissue. This sequence represents a human CA gene of the invention. Note:
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences
SQ	
Sequence	76180 BP; 17603 A; 18463 C; 19657 G; 20457 T; 0 U; 0 Other;
Query Match	6.5%; Score 36.4; DB 13; Length 76180;
Best Local Similarity	59.8%; Pred. No. 41;
Matches	61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
OY	25 GGACGACCACTCAGCGTGACTCACCCCATGTGGCTTGAAGCGAAGGCCTTCGA 84 75766 GGGGGCAGCCCTGTCATGTGTACTTTGGCAGATTGAGAGAACAACAGAGCAGCGAG 75707
OY	85 GGCAGGGCCAGGCGACGCCCTCAGGTGGGTGACGGCAGGGGTC 126 75706 GGAGGGAGAGCGCAGCCACCAGGAGGTGACAGGTGACATC 75665
Dt	
Db	
RESULT 13	
ID AD251725	
ID AD251725	standard; DNA; 138203 BP.
XX AC	
XX AD251725;	
DT DT	14-JUL-2005 (first entry)
DX DX	FR-008 polypeptide gene cluster, SEQ ID 1.
KM KM	Antifungal; antiparasitic; polypeptide; gene; ds.
OS Streptomyces sp.; FR-008.	
FH FH	
Key Location/Qualifiers	
CDS CDS	/caga= a
FT FT	complement(574..1950)
FT FT	/caga= b
CDS CDS	complement("PAD-dependent monooxygenase, Paco, SEQ ID 20"
FT FT	complement(2264..3037)
FT FT	/*caga= b
CDs CDs	product="ADC lyase, PabC, SEQ ID 22"
CDs CDs	complement(3150..3818)

FT	/tag= c
FT	/product= "Regulator protein, FecRI, SEQ ID 10"
FT	/complement(4377. .7205)
FT	/tag= d
FT	/product= "Regulator protein, FecRIT, SEQ ID 11"
FT	/complement(7210. .10320)
FT	/tag= e
FT	/product= "Regulator protein, FecRIII, SEQ ID 12"
FT	/complement(10298. .13315)
FT	/tag= f
FT	/product= "Regulator protein, FecRIV, SEQ ID 13"
FT	/complement(14698
FT	/tag= g
FT	/product= "Glycosyltransferase, FecMI, SEQ ID 17"
FT	/complement(14953. .16011
FT	/tag= h
FT	/product= "GDP-ketosugar aminotransferase, FecMIT, SEQ ID
FT	18"
FT	/complement(16106. .17287
FT	/tag= i
FT	/product= "Cytochrome P-450 monooxygenase, FecP, SEQ ID
FT	14"
FT	/complement(17334. .17528
FT	/tag= j
FT	/product= "Ferredoxin protein, FecFE, SEQ ID 15"
FT	/complement(17556. .18413
FT	/tag= k
FT	/product= "Thioesterase, FecTE, SEQ ID 16"
FT	/complement(18610. .20781
FT	/tag= l
FT	/product= "ADC synthase, PadAB, SEQ ID 21"
FT	/complement(20927. .26158
FT	/tag= m
FT	/product= "PKS FecA, SEQ ID 2"
FT	/complement(26333. .27340
FT	/tag= n
FT	/product= "ABC transporter, FecTI, SEQ ID 8"
FT	/complement(27561. .28280
FT	/tag= o
FT	/product= "ABC transporter, FecTII, SEQ ID 9"
FT	/complement(28983. .60860)
FT	/tag= p
FT	/product= "PKS FecC, SEQ ID 3"
FT	/complement(60962. .77587)
FT	/tag= q
FT	/product= "PKS FecB, SEQ ID 4"
FT	/complement(77983. .84132)
FT	/tag= r
FT	/product= "PKS FecF, SEQ ID 5"
FT	/complement(84170. .107485)
FT	/tag= s
FT	/product= "PKS FecE, SEQ ID 6"
FT	/complement(107496. .136148)
FT	/tag= t
FT	/product= "PKS Fecd, SEQ ID 7"
FT	/complement(136558. .137766)
FT	/tag= u
FT	/product= "GDP-mannose-4,6-dehydratase, FecMTIII, SEQ ID
FT	19"
XX	
XX	US2005089884-A1.
XX	
XX	28-APR-2005.
XX	
XX	06-APR-2004; 2004US-00819386.
XX	
XX	23-OCT-2003; 2003KR-00074035.
XX	
XX	(KOAD) KOREA ADV INST SCI & TECHNOLOGY.
XX	(UTSH-) UNIV SHANGHAI JIAOTONG.
XX	
XX	Lee SY, Deng Z, Chen S, Jeong KJ, Zhou X;
XX	

XX Human ALD gene exon-3+4 region.
DB ALD gene; adrenoleukodystrophy; adrenomyeloneuropathy; diagnosis;
KW gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 1
FT /tag= a
FT /note= "base n at position 1 represents the 5' sequence
FT of intron-2, not given in the specification"
FT 2..151
FT /tag= b
FT /label= Intron-2
FT /note= "150 bases of the intron-2 3' sequence"
FT 152..294
FT /tag= c
FT /label= Exon-3
FT 295..384
FT /tag= d
FT /label= Intron-3
FT 385..553
FT /tag= d
FT /label= Exon-4
FT 554..703
FT /tag= e
FT /label= Intron-4
FT /note= "150 bases of the intron-4 5' sequence"
FT 704
FT /tag= f
FT /note= "base n at position 704 represents approx. 350 bp
FT of intron-4 not given in the specification"
FT
FT
XX CA2108606-A.
XX
XX 16-APR-1995.
XX
XX 15-OCT-1993; 93CA-02108606.
XX
XX 15-OCT-1993; 93CA-02108606.
XX
XX 15-OCT-1993; 93CA-02108606.
XX
XX (INRM) INST NAT SANTE & RECH MEDICALE.
XX
XX
XX Mandel J, Aubourg P, Mosser J, Sarde C;
XX
XX WPI; 1995-215721/29.
XX
XX
XX New nucleic acid responsible for adreno-leuco-dystrophy - related probes,
XX proteins and antibodies, useful for diagnosis and treatment.
XX
XX Claim 3; Fig 7B; 39pp; English.
XX
XX
XX The putative ALD gene identified in the distal part of chromosome Xq28
XX had deletions in 1 or several exons in 6 of 85 independent ALD patients
XX examined. The gene has 10 exons and 9 introns, shown in AA094042-47
XX
XX
XX Sequence 704 BP; 127 A; 235 C; 205 G; 135 T; 0 U; 2 Other;
SQ
Query Match 6.3%; Score 35.8; DB 2; Length 704;
Best Local Similarity 51.6%; Pred. No. 17; Mismatches 77; Indels 0; Gaps 0;
Matches 82; Conservative 0;
OY 386 CTGAAGGCGCATGCTTGAATTAGTAGAGGGTGGGCTGGGCGAGGCCACAG 445
DB 670 CAGAAAGCACATGAGAGTCCCTGATGAGGAGCGGGGATAGAGAGCTGGGGCAGCAG 611
OY 446 TCTGAGTCAGAGCCGAGAGCGAGAGCTGCTCCAGACATGCGCCGCTTGGCAGT 505
DB 610 TCAGACCTGACAGCAGCAGCGAGCGGGGCTCATCATAGGAGCGGACAGCTTACCTC 551
OY 506 CAGTCCTCTGCGCCACTGAGAACAGCTGTAGAGAGC 544

DB 550 GGATCTTCAGGGGGCCCTTCACACGACACCAAGCCGCGC 512

Search completed: March 25, 2006, 16:28:07
Job time : 476 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 13:42:31 ; Search time 959 seconds
(without alignments)
1304.021 Million cell updates/sec

Title: US-10-798-652-6

Perfect score: 22
Sequence: 1 cttccctcaccctccagcacatg 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 1641224

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_ses:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hrg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	15.6	70.9	25	6	AX701087 Sequence
c 2	15.2	69.1	22	6	BD086268 Vector. 8
c 3	15.2	69.1	22	6	AX003404 Sequence
c 4	14.6	66.4	24	6	AS7518 Sequence 10
c 5	14.6	66.4	24	6	AR052984 Sequence
c 6	14.6	66.4	30	6	AR306619 Sequence
c 7	14.6	66.4	30	6	AR306619 Sequence
c 8	14.6	66.4	30	6	AR340054 Sequence
c 9	14.6	66.4	30	6	AR340063 Sequence
c 10	14.6	66.4	30	6	AR412123 Sequence
c 11	14.6	66.4	30	6	AR412132 Sequence
c 12	14.4	65.5	21	6	BD174200 Sequence
c 13	14.4	65.5	21	6	BD185155 Sequence
c 14	14.4	65.5	22	6	AX817728 Sequence
c 15	14.2	64.5	24	6	AS7517 Sequence
c 16	14.2	64.5	24	6	AR052983 Sequence
c 17	14.2	64.5	24	6	AR096578 Sequence
c 18	14.2	64.5	24	6	AR105954 Sequence

19	14.2	64.5	24	6	AR442798 Sequence
c 20	14	63.6	24	6	AR567736 Sequence
c 21	13.8	62.7	20	6	BD210000 Nucleic a
c 22	13.8	62.7	20	6	AR654868 Sequence
c 23	13.8	62.7	20	6	AX010496 Sequence
c 24	13.8	62.7	24	6	AX958559 Sequence
c 25	13.8	62.7	24	6	AX958815 Sequence
c 26	13.6	61.8	20	6	AR25902 Sequence
c 27	13.6	61.8	20	6	AR373824 Sequence
c 28	13.6	61.8	20	6	AX418821 Sequence
c 29	13.6	61.8	22	6	B51058 Sequence
c 30	13.6	61.8	24	6	AS7510 Sequence 2
c 31	13.6	61.8	24	6	AS7512 Sequence 4
c 32	13.6	61.8	24	6	AR052976 Sequence
c 33	13.6	61.8	24	6	AR052978 Sequence
c 34	13.6	61.8	25	6	CO862052 Sequence
c 35	13.6	61.8	27	6	AR091427 Sequence
c 36	13.6	61.8	27	6	AR091430 Sequence
c 37	13.6	61.8	27	6	AR125632 Sequence
c 38	13.6	61.8	27	6	AR125635 Sequence
c 39	13.4	60.9	25	6	AX610213 Sequence
c 40	13.2	60.0	19	6	AR153219 Sequence
c 41	13.2	60.0	19	6	AR163234 Sequence
c 42	13.2	60.0	19	6	AR176959 Sequence
c 43	13.2	60.0	19	6	AR210978 Sequence
c 44	13.2	60.0	19	6	AR431659 Sequence
c 45	13.2	60.0	19	6	AR533196 Sequence

ALIGNMENTS

RESULT 1
LOCUS AX701087 25 bp DNA linear PAT 03-APR-2003
DEFINITION Sequence 86 from Patent WO03011334.
ACCESSION AX701087
VERSION AX701087.1 GI:29536869
KEYWORDS
SOURCE
ORGANISM
other sequences; artificial sequences.
REFERENCE
1
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
ORIGIN
Query Match 70.9%; Score 15.6; DB 6; Length 25;
Best Local Similarity 81.8%; Pred. No. 2.6e+04;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Ct 1 cttccctcaccctccagcacatg 22
Db 22 cttccctcaccctccagcacatg 1
RESULT 2
LOCUS BD086268 22 bp DNA linear PAT 27-AUG-2002
DEFINITION Vector.
ACCESSION BD086268.1 GI:22631878
VERSION JP 2001525168-A/21.
KEYWORDS
SOURCE
synthetic construct

ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 22)
AUTHORS Shen,S., Schedl,A. and Harmar,A.J.
TITLE Vector
JOURNAL Patent: JP 2001525168-A 21 11-DEC-2001;
MEDICAL RESEARCH COUNCIL
OS Artificial Sequence
PN JP 2001525168-A/21
PD 11-DEC-2001
PF 27-NOV-1998 JP 2000523326
PR 28-NOV-1997 GB 9725311.6,28-NOV-1997 GB 9725313.2 PR
20-MAR-1998 GB 9806072.6,05-NOV-1998 GB 9824275.3 PI
SANBING SHEN,ANDREAS SCHEDL,ANTHONY JOHN HARMAR PC
C12N15/09, C12N15/00
CC Description of Artificial Sequence: Primer
FT Key Location/Qualifiers
FT source 1..22 /organism='Artificial Sequence'.
FEATURES
source 1..22
Location/Qualifiers
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 69.1%; Score 15.2; DB 6; Length 22;
Best Local Similarity 85.0%; Pred.No.4e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TCCCTCAGCTCCAGCCACAT 21
1 TTCATCAGCTCCATCCACAT 20

Db 1 TTCATCAGCTCCATCCACAT 20

RESULT 3
AX003404 22 bp DNA linear PAT 24-AUG-2000
LOCUS AX003404
DEFINITION Sequence 21 from Patent WO928449.
ACCESSION AX003404
VERSION AX003404.1 GI:9927208
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE 1
AUTHORS Schedl,A. and Harmar,A.J.
TITLE Vectors
JOURNAL Patent: WO 928449-A 21 10-JUN-1999;
SCHIDL ANDREAS (DE); HARMAR ANTHONY JOHN (GB)
FEATURES
source 1..22
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Primer"

ORIGIN
Query Match 69.1%; Score 15.2; DB 6; Length 22;
Best Local Similarity 85.0%; Pred.No.4e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TCCCTCAGCTCCAGCCACAT 21
1 TTCATCAGCTCCATCCACAT 20

Db 1 TTCATCAGCTCCATCCACAT 20

RESULT 4
A57518 24 bp DNA linear PAT 03-MAR-1998
LOCUS A57518
DEFINITION Sequence 10 from Patent WO9632483.
ACCESSION A57518
VERSION A57518.1 GI:3713376

KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.

REFERENCE 1
AUTHORS Masucci,M.G.
TITLE IMMUNE-EVADING PROTEINS
JOURNAL Patent: WO 9632483-A 10 17-OCT-1996;
MASUCCI MARIA GRAZIA (SE)
COMMENT Other publication AU 5284296 961030.
FEATURES
source 1..24
Location/Qualifiers
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 66.4%; Score 14.6; DB 6; Length 24;
Best Local Similarity 81.0%; Pred.No.7.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCACAT 21
4 CACCCGACCTCCAGCTCCAT 24

Db 4 CACCCGACCTCCAGCTCCAT 24

RESULT 5
AR052984 24 bp DNA linear PAT 29-SEP-1999
LOCUS AR052984
DEFINITION Sequence 16 from patent US 5833991.
ACCESSION AR052984
VERSION AR052984.1 GI:5977846
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
unclassified.

REFERENCE 1 (bases 1 to 24)
AUTHORS Masucci,M.G.
TITLE Glycine-containing sequences conferring inviability to the immune system
JOURNAL Patent: US 5833991-A 16 10-NOV-1998;
FEATURES
source 1..24
Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 66.4%; Score 14.6; DB 6; Length 24;
Best Local Similarity 81.0%; Pred.No.7.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCACAT 21
4 CACCCGACCTCCAGCTCCAT 24

Db 4 CACCCGACCTCCAGCTCCAT 24

RESULT 6
AR306619 30 bp DNA linear PAT 12-JUN-2003
LOCUS AR306619
DEFINITION Sequence 10 from patent US 6548642.
ACCESSION AR306619
VERSION AR306619.1 GI:31696821
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
unclassified.

REFERENCE 1 (bases 1 to 30)
AUTHORS Kieliszewski,M.J.
TITLE Synthetic genes for plant gums
JOURNAL Patent: US 6548642-A 10 15-APR-2003;
Ohio University; Athens, OH
FEATURES
source 1..30
Location/Qualifiers

/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 7.4e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 5 CACCTTCACCTCCACCCCAT 25

RESULT 7
AR306628 30 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 38 from patent US 6548642.
DEFINITION AR306628
ACCESSION AR306628.1 GI:31696830
VERSION
KEYWORDS
SOURCE
ORGANISM
Unclasseified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Kieliszewski,M.J.
TITLE Synthetic genes for plant gums
JOURNAL Patent: US 6548642-A 38 15-APR-2003;
Ohio University; Athens, OH
FEATURES
source 1..30
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 7.4e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 5 CACCTTCACCTCCACCCCAT 25

RESULT 8
AR340054 30 bp DNA linear PAT 17-AUG-2003
LOCUS Sequence 10 from patent US 6570062.
DEFINITION AR340054
ACCESSION AR340054.1 GI:33731348
VERSION
KEYWORDS
SOURCE
ORGANISM
Unclasseified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Kieliszewski,M.J.
TITLE Synthetic genes for plant gums and other hydroxyproline-rich
glycoproteins
JOURNAL Patent: US 6570062-A 10 27-MAY-2003;
Ohio University; Athens, OH
FEATURES
source 1..30
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 7.4e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 5 CACCTTCACCTCCACCCCAT 25

RESULT 9
AR340063 30 bp DNA linear PAT 17-AUG-2003
LOCUS Sequence 38 from patent US 6570062.
DEFINITION AR340063
ACCESSION AR340063.1 GI:33731357
VERSION
KEYWORDS
SOURCE
ORGANISM
Unclasseified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Kieliszewski,M.J.
TITLE Synthetic genes for plant gums and other hydroxyproline-rich
glycoproteins
JOURNAL Patent: US 6570062-A 38 27-MAY-2003;
Ohio University; Athens, OH
FEATURES
source 1..30
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 7.4e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 5 CACCTTCACCTCCACCCCAT 25

RESULT 10
AR412123 30 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 10 from patent US 6639050.
DEFINITION AR412123
ACCESSION AR412123.1 GI:40166767
VERSION
KEYWORDS
SOURCE
ORGANISM
Unclasseified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Kieliszewski,M.J.
TITLE Synthetic genes for plant gums and other hydroxyproline-rich
glycoproteins
JOURNAL Patent: US 6639050-A 10 28-OCT-2003;
Ohio University; Athens, OH
FEATURES
source 1..30
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 7.4e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 5 CACCTTCACCTCCACCCCAT 25

RESULT 11
AR412132 30 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 38 from patent US 6639050.
DEFINITION AR412132
ACCESSION AR412132.1 GI:40166776
VERSION
KEYWORDS
SOURCE
ORGANISM
Unclasseified.

REFERENCE	1 (bases 1 to 30)
AUTHORS	Kieliszewski, M.J.
TITLE	Synthetic genes for plant gums and other hydroxyproline-rich glycoproteins
JOURNAL	Patent: US 6639050-A 38 28-OCT-2003; Ohio University; Athens, OH
FEATURES	location/Qualifiers
source	1..30 /organism="unknown" /mol_type="genomic DNA"
ORIGIN	
Query Match	66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity	81.0%; Pred. No. 7.4e+04;
Matches	17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	1 CTCCTCACCTCCAGCCACAT 21
DB	5 CACCTTCACCTCCACCCCAT 25
RESULT 12	
BD174200/c	21 bp DNA linear PAT 18-FEB-2003
LOCUS	BD174200
DEFINITION	Periplastic converting agent.
ACCESSION	BD174200
VERSION	BD174200.1 GI:28415537
KEYWORDS	WO 02066049-A/46.
SOURCE	synthetic construct
ORGANISM	synthetic construct other sequences: artificial sequences.
REFERENCE	1 (bases 1 to 21)
AUTHORS	Hikichi, Y., Shintani, Y. and Matsui, H.
TITLE	Periplastic converting agent
JOURNAL	Patent: WO 02066049-A 46 23-AUG-2002; TAKEDA CHEMICAL INDUSTRIES LTD, YUKIKO HIKICHI, YASUSHI SHINTANI, HIDEKI MATSUI
COMMENT	OS Artificial Sequence PN WO 02066049-A/46 PD 23-AUG-2002 PF 21-FEB-2002 WO 2002JP001536 PR 23-FEB-2001 JP 01P 09450 PI YUKIKO HIKICHI, YASUSHI SHINTANI, HIDEKI MATSUI PC A61K38/17, A61K31/711, A61K48/00, A61P43/00, A61P21/00, A61P21/04, PC A61P15/00. PC C12N15/12, C07K14/47 CC Primer FH Key location/Qualifiers FT source 1..21 FT /organism='Artificial Sequence'.
FEATURES	Location/Qualifiers
source	1..21 /organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"
ORIGIN	
Query Match	65.5%; Score 14.4; DB 6; Length 21;
Best Local Similarity	93.8%; Pred. No. 9.4e+04;
Matches	15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 CTCCTCACCTCCAGC 16
DB	18 CTCCTCACCTCCATC 3
RESULT 13	
BD185155/c	21 bp DNA linear PAT 17-JUN-2003
LOCUS	BD185155
DEFINITION	Cell differentiating agent.
ACCESSION	BD185155
VERSION	BD185155.1 GI:31877355
KEYWORDS	WP 2002356438-A/46.

SOURCE	synthetic construct
ORGANISM	synthetic construct
REFERENCE	other sequences; artificial sequences.
AUTHORS	1 (bases 1 to 21)
TITLE	Hikichi,Y., Shintani,Y. and Matsui,H.
JOURNAL	Cell differentiating agent Patent: JP 2002356438-A 46 13-DEC-2002;
COMMENT	TAKEDA CHEMICAL INDUSTRIES LTD OS Artificial Sequence PN JP 2002356438-A/46 PD 13-DEC-2002 PF 21-FEB-2002 JP 2002044741 PI YUKIKO HIKICHI,YASUSHI SHINTANI,HIDEKI MATSUI PC A61K38/00,A61K31/7088,A61P15/00,A61P21/04,A61P35/00,C12N15/09// PC C07K14/525 PC C12Q1/68,A61K37/02,C12N15/00 CC Primer FH Key FT source FT Location/Qualifiers 1..21 /organism='Artificial Sequence'.
FEATURES	source 1..21 Location/Qualifiers /organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"
ORIGIN	Query Match 65.5%; Score 14.4; DB 6; Length 21; Best Local Similarity 93.8%; Pred. No. 9.4e+04; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY	1 CTCCTCATCCTCCAGC 16 18 CTCCTCATCCTCCATC 3
Db	
RESULT 14	
LOCUS	AX817728 22 bp DNA linear PAT 10-DEC-2003
DEFINITION	Sequence 476 from Patent WO02081517.
ACCESSION	AX817728
VERSION	AX817728.1 GI:39722920
KEYWORDS	.
SOURCE	synthetic construct synthetic construct other sequences; artificial sequences.
REFERENCE	1
AUTHORS	Deriastofaro,M.F., Padigaru,M., Miller,C., Tchernen,V., Zhong,H., Zhong,M., Anderson,D., Ballinger,R., Gerlach,V., Spytek,K.A., Rastelli,L., Kerkuta,R., Guo,X., Zernusen,B., Andrew,D., Meres,P., Pactunajin,M., Burgess,C.E., Eisen,A., Wolenc,A., Baumgartner,J., Shimkets,R.A., Gusev,V., Vernet,C.A., Taupier,R.J., Pena,C., Shenoy,S., Li,L., Casman,S., Boljog,F., Fernandes,E., Smithson,G., Malvankar,U., Tallion,B. and Liu,X. Novel polypeptides and nucleic acids encoded thereby Patent: WO 02081517-A 476 17-OCT-2002; Curagen Corporation (US) Location/Qualifiers 1..22 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="Description of Artificial Sequence: PCR Primer sequence"
FEATURES	source
TITLE	
JOURNAL	
ORIGIN	Query Match 65.5%; Score 14.4; DB 6; Length 22; Best Local Similarity 93.8%; Pred. No. 9.4e+04; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY	1 CTCCTCATCCTCCAGC 16 1 CTCCTCATCCTCCATC 16
Db	

RESULT 15
A57517/c

A57517/c

LOCUS	A57517	24 bp	DNA	linear	PAT 03-MAR-1998
DEFINITION	Sequence 9 from Patent WO9632483.				

ACCESSION A57517

ACCESSION A57517

ACCESSION A57517

KEYWORDS .

SOURCE unidentified

ORGANISM unidentified

unclassified sequences.

REFERENCE
1
Masucci M C

TITLE IMMUNE-EVADING PROTEINS

JOURNAL Patent: WO 9632483-A 9

MASUCCI MARIA GRAZIA (S)

COMMENT Other publication AU 5284296 961030.

FEATURES

Location/Qualifiers

Source

```
/organism="unidentified"
```

```
/mot_cype=unassigned DNA
/dh_yref="rayon:3264A"
```

1203010000 - 7070 - 000 /

ORIGIN

Query match 64.5%; Score 14.2; DB 6; Length 24;

Best Local Similarity 84.2%; Pred. No. 1.2e+05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

3 CCCTGACCTCCAGGCACAT 21

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

Search completed: March 25, 2006, 14:43:29
Job time : 962 secs

Job time : 962 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 13:22:55 / Search time 372.5 Seconds
(without alignments)
393.619 Million cell updates/sec

Title: US-10-798-652-6

Perfect score: 22
Sequence: 1 ctccctcacctccagcacatg 22

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 4138570

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*

1:	geneseqn1980s:*
2:	geneseqn1990s:*
3:	geneseqn2000s:*
4:	geneseqn2001s:*
5:	geneseqn2001bs:*
6:	geneseqn2002as:*
7:	geneseqn2002bs:*
8:	geneseqn2003as:*
9:	geneseqn2003bs:*
10:	geneseqn2003cs:*
11:	geneseqn2003ds:*
12:	geneseqn2004as:*
13:	geneseqn2004bs:*
14:	geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	13	ADT97906 Human FGF
2	20.4	97.7	22	13	ADT97907 Human FGF
3	15.6	70.9	25	10	ADB67764 HIV-1 gp1
4	15.2	69.1	22	2	AAK57861 PCR prime
5	15.2	69.1	22	5	AAK85242 Reverse P
6	15.2	69.1	25	9	ACH62883 DNA large
7	15.2	69.1	25	9	ACH63009 DNA large
8	14.6	66.4	24	2	AAV55821 Multimeri
9	14.6	66.4	29	14	ADY00944 PCR prime
10	14.6	66.4	30	2	AAK27663 DNA encod
11	14.6	66.4	30	6	ABL51730 HRGP rela
12	14.6	66.4	30	6	ABL51739 Hydroxypr
13	14.6	66.4	30	13	ADU09143 Ser-Hyd4
14	14.6	66.4	30	13	ADU09116 Hydroxypr
15	14.4	65.5	21	6	AA149623 Tumour di
16	14.4	65.5	21	12	ADM69088 Human TAT
17	14.4	65.5	21	13	ADG74003 A. gossyp
18	14.4	65.5	22	8	ABT33560 NOV forwa
19	14.2	64.5	24	2	AAQ40923 Probe Pol

C	20	14.2	64.5	24	2	AAT39970 Minimal m
C	21	14.2	64.5	24	2	AAV55820 Multimeri
C	22	14.2	64.5	28	10	AB224297 Human MIS
C	23	14.2	64.5	30	4	AAH21949 Mouse dig
C	24	14.2	64.5	30	12	ADK15423 PCR prime
C	25	14.2	64.5	30	12	ADQ62731 Human int
C	26	14	63.6	21	13	ADU43700 Knock-dow
C	27	14	63.6	24	6	ABK14322 Human pzy
C	28	14	63.6	25	9	AC149317 Human mic
C	29	13.8	62.7	17	11	ADL47245 Human NOG
C	30	13.8	62.7	19	13	ADK16045 Human DNA
C	31	13.8	62.7	20	3	AAZ24490 wheat sol
C	32	13.8	62.7	20	14	AAE51146 pyranalR
C	33	13.8	62.7	21	10	ADG30065 IL2-targe
C	34	13.8	62.7	21	10	ADG30061 IL2-targe
C	35	13.8	62.7	21	10	ADG30057 IL2-targe
C	36	13.8	62.7	21	12	ADQ44286 Swine Mx1
C	37	13.8	62.7	23	10	ADG28531 IL2 g1NA-
C	38	13.8	62.7	23	10	ADG30053 IL2-targe
C	39	13.8	62.7	24	12	ADP32021 Probe #3
C	40	13.8	62.7	25	2	AAK81735 PCR prime
C	41	13.8	62.7	27	14	ADZ85335 Interleuk
C	42	13.6	61.8	20	4	AAK07411 PCR prime
C	43	13.6	61.8	20	4	AAK12159 Rat PTP1B
C	44	13.6	61.8	20	6	ABK85234 Rat PTP1B
C	45	13.6	61.8	20	6	ABK37403 Rat PTP1B

ALIGNMENTS

RESULT 1	ADT97906	ADT97906 standard; DNA; 22 BP.
ID	ADT97906	
XX	ADT97906;	
AC	27-JAN-2005 (first entry)	
XX		
DT	Human FGF-3 promoter 5' proximal region T allele probe.	
XX		
DE	Human; ss; fibroblast growth factor-3; FGF-3; cancer; oesophageal cancer;	
XX	breast cancer; ovarian cancer; prostate cancer; head and neck cancer;	
KW	oesophageal squamous cell carcinoma; PCR; probe; SNP;	
KW	single nucleotide polymorphism.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	modified_base 1	/*tag= a
FT	modified_base 22	/mod_base= OTHER
FT	modified_base 22	/note= "modified with vic (not defined)"
FT		/*tag= b
FT		/mod_base= OTHER
FT		/note= "g is covalently linked to a TAMRA moiety (6-carbotetramethyl-rhodamine)"
XX	US2004219582-A1.	
XX	04-NOV-2004.	
XX	11-MAR-2004; 2004US-00798652.	
XX	17-MAR-2003; 2003US-0455689P.	
XX	(GUOY/) GUO Y.	
XX	Guo Y;	
XX	WPI; 2004-794435/78.	
XX		

PT Novel isolated nucleic acid molecule having single nucleotide
PT polymorphism in upstream untranslated region of fibroblast growth factor-
PT 3 gene, useful for assessing related susceptibility of mammal to cancer.
XX
XX
PS Claim 34; SEQ ID NO 6; 15bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising the
CC upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3
CC gene sequence (ADT97901) with a single nucleotide polymorphism (SNP) at
CC position 69. Also included are an isolated nucleic acid molecule
CC comprising a sequence complementary to ADT97901, a vector comprising
CC ADT97901 operably linked to a reporter gene, a host cell comprising the
CC vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating
CC a nucleic acid sample from the mammal, and determining whether a cytosine
CC or thymine is present at position 69 of ADT97901), a kit for performing
CC the method (comprising a first oligonucleotide probe which anneals
CC specifically with the target portion of the mammal's genome, where the
CC first probe comprises a first fluorescent label and a first fluorescence
CC quencher attached to its separate nucleotide residues and the target
CC portion includes the nucleotide residue located at position 69 of
CC ADT97901, and a pair of primers for amplifying a reference portion of the
CC FGF-3 gene, where the reference portion includes the nucleotide residue
CC located at position 69 of ADT97901) and a microarray having at least one
CC oligonucleotide probe that can anneal with a target portion of a mammal's
CC genome, where the target portion includes the nucleotide residue located
CC at position 69 of ADT97901. The method is useful for detecting SNP in FGF
CC -3 gene in a mammal, preferably a human, and is also useful for assessing
CC the relative susceptibility of a mammal to cancer (especially oesophageal
CC cancer), which shows an association with the presence of the C-allele.
CC The cancer is chosen from oesophageal, breast, ovarian, prostate, head
CC and neck cancer. The oesophageal cancer is oesophageal squamous cell
CC carcinoma. The present sequence is a Tagman PCR probe used to genotype
CC individuals for the T SNP in the UTR of FGF-3.
SQ
SQ Sequence 22 BP; 4 A; 12 C; 2 G; 4 T; 0 U; 0 Other;
Query Match 100.0%; Score 22; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CTCCTCACCCTCCAGCCACATG 22
DB 1 CTCCTCACCCTCCAGCCACATG 22
RESULT 2
ADT97907
ID ADT97907 standard; DNA; 22 BP.
XX
XX ADT97907;
XX
DT 27-JUN-2005 (first entry)
XX
XX Human FGF-3 promoter 5' proximal region C allele probe.
DE
XX
XX Human; ss: fibroblast growth factor-3; FGF-3; cancer; oesophageal cancer;
KW breast cancer; ovarian cancer; prostate cancer; head and neck cancer;
KW oesophageal squamous cell carcinoma; PCR; probe; SNP;
KW single nucleotide polymorphism.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH modified_base 1
FT /*tag= a
FT /mod_base= OTHER
FT /note= "C is covalently linked to a 6-carbo fluorescein
FT moiety (FAM)"
FT 22
FT modified_base
FT /*tag= b
FT /mod_base= OTHER
FT /note= "G is covalently linked to a TAMRA moiety (6-
FT carbocetramethyl-rhodamine)"

XX
XX US2004219582-A1.
XX
XX 04-NOV-2004.
XX
XX
XX 11-MAR-2004; 2004US-00798652.
XX
XX 17-MAR-2003; 2003US-0455689P.
XX
XX (GUOY/) GUO Y.
XX
XX Guo Y;
XX
XX WPI; 2004-794435/78.
XX
XX
XX Novel isolated nucleic acid molecule having single nucleotide
PT polymorphism in upstream untranslated region of fibroblast growth factor-
PT 3 gene, useful for assessing related susceptibility of mammal to cancer.
XX
XX
PS Claim 20; SEQ ID NO 7; 15bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising the
CC upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3
CC gene sequence (ADT97901) with a single nucleotide polymorphism (SNP) at
CC position 69. Also included are an isolated nucleic acid molecule
CC comprising a sequence complementary to ADT97901, a vector comprising
CC ADT97901 operably linked to a reporter gene, a host cell comprising the
CC vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating
CC a nucleic acid sample from the mammal, and determining whether a cytosine
CC or thymine is present at position 69 of ADT97901), a kit for performing
CC the method (comprising a first oligonucleotide probe which anneals
CC specifically with the target portion of the mammal's genome, where the
CC first probe comprises a first fluorescent label and a first fluorescence
CC quencher attached to its separate nucleotide residues and the target
CC portion includes the nucleotide residue located at position 69 of
CC ADT97901, and a pair of primers for amplifying a reference portion of the
CC FGF-3 gene, where the reference portion includes the nucleotide residue
CC located at position 69 of ADT97901) and a microarray having at least one
CC oligonucleotide probe that can anneal with a target portion of a mammal's
CC genome, where the target portion includes the nucleotide residue located
CC at position 69 of ADT97901. The method is useful for detecting SNP in FGF
CC -3 gene in a mammal, preferably a human, and is also useful for assessing
CC the relative susceptibility of a mammal to cancer (especially oesophageal
CC cancer), which shows an association with the presence of the C-allele.
CC The cancer is chosen from oesophageal, breast, ovarian, prostate, head
CC and neck cancer. The oesophageal cancer is oesophageal squamous cell
CC carcinoma. The present sequence is a Tagman PCR probe used to genotype
CC individuals for the C SNP in the UTR of FGF-3.
XX
XX
SQ Sequence 22 BP; 3 A; 12 C; 3 G; 4 T; 0 U; 0 Other;
Query Match 92.7%; Score 20.4; DB 13; Length 22;
Best Local Similarity 95.5%; Pred. No. 11e+02;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CTCCTCACCCTCCAGCCACATG 22
DB 1 CTCCTCACCCTCCAGCCACATG 22
RESULT 3
ADB67764/C
ID ADB67764 standard; DNA; 25 BP.
XX
XX ADB67764;
XX
XX 04-DEC-2003 (first entry)
XX
XX HIV-1 gp120 optimised sequencing primer g120c-SF4.
DE
XX
XX HIV; Tat; HIV-1; Nef; gp120; vaccine; immunisation; human; anti-HIV;
KW primer; ss.
XX

OS Synthetic.
 OS Human immunodeficiency virus 1.
 XX
 PN WO2003011334-A1.
 XX
 PD 13-FEB-2003.
 XX
 PF 26-JUL-2002; 2002WO-EP008343.
 XX
 PR 27-JUL-2001; 2001GB-00018367.
 XX
 PA (GLAX) GLAXOSMITHKLINE BIOLOGICALS SA.
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Excl PF, Tlce JP, Van Wely CA, Voss G;
 XX
 DR WPI; 2003-239474/23.
 XX
 PT Use of an HIV Tat, Nef or gp120 protein or polynucleotide in
 PT manufacturing a vaccine for a prime-boost delivery for the prophylactic
 PT or therapeutic immunization of humans against HIV.
 XX
 PS Example 15; Fig 21; 108pp; English.
 XX
 CC The present invention describes the use of an HIV Tat protein or
 CC polynucleotide, an HIV Nef protein or polynucleotide, or an HIV Tat
 CC protein or polynucleotide linked to an HIV Nef protein or polynucleotide,
 CC and an HIV gp120 protein or polynucleotide in manufacturing a vaccine for
 CC a prime-boost delivery for the prophylactic or therapeutic immunisation
 CC of humans against HIV. The protein or polynucleotide is delivered via a
 CC bombardment approach. Also described: (1) a recombinant DNA molecule
 CC comprising a Nef and/or Tat and/or gp120 gene in a vector in which the
 CC gene of interest is inserted 3' to an enhanced HCMV IE1 promoter; (2)
 CC particles, preferably gold particles, coated with recombinant DNA
 CC comprising a Nef and/or Tat and/or gp120 gene in a vector; and (3) a kit
 CC comprising at least two different vaccine compositions including: (a) a
 CC composition comprising particles coated with DNA encoding gp120 and nef
 CC and/or tat or nef/tat; and (b) a composition comprising gp120 and nef
 CC and/or tat or nef/tat DNA or proteins, where the DNA or proteins are not
 CC coated onto the particles. The HIV proteins or polynucleotides are used
 CC to produce vaccines having anti-HIV activity. The recombinant DNA
 CC molecule can be used in manufacturing an HIV vaccine for the prophylactic
 CC or therapeutic immunisation of humans. The present sequence for the represents an
 CC oligonucleotide primer for sequencing optimised gp120, which is used in
 CC an example from the present invention.
 XX
 SQ Sequence 25 BP; 6 A; 4 C; 12 G; 3 T; 0 U; 0 Other;
 Query Match 70.9%; Score 15.6; DB 10; Length 25;
 Best Local Similarity 81.8%; Pred. No. 9.1e+03;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CTCCTCAGCTTCAGCCACATG 22
 DB 22 CTTCCACACCTCCGCGCACATG 1
 RESULT 4
 AAX57861
 ID AAX57861 standard; DNA; 22 BP.
 XX
 AC AAX57861;
 XX
 DT 15-JUL-1999 (first entry)
 XX
 DE PCR primer used in construction of yeast artificial chromosome.
 XX
 KM YAC; yeast artificial chromosome; PCR primer; sexual dysfunction;
 KM reporter gene; transgenic mammal; therapy; circadian function;
 KM sleep disorder; eating disorder; premenstrual syndrome; birth defect;
 KM autoimmune disorder; ss.
 XX
 OS Synthetic.

XX
 XX GB2331752-A.
 XX
 PN 02-JUN-1999.
 XX
 PD 27-NOV-1998; 98GB-00026126.
 XX
 PF 28-NOV-1997; 97GB-00025311.
 XX
 PR 28-NOV-1997; 97GB-00025313.
 PR 20-MAR-1998; 98GB-00006072.
 PR 05-NOV-1998; 98GB-00024275.
 XX
 PA (MED1-) MEDICAL RES COUNCIL.
 PA
 PI Shen S, Schedl A, Harmar AJ;
 XX
 DR WPI; 1999-290603/25.
 XX
 PT New reporter gene labeled YAC vectors and transgenic mammals used for
 PT screening potential active agents.
 XX
 PS Disclosure; Page 56; 98pp; English.
 XX
 CC This sequence represents a PCR primer used in the construction of a yeast
 CC artificial chromosome of the invention. The yeast artificial chromosome
 CC (YAC) vectors contain a reporter gene and transgenic mammals produced
 CC using them may be used to screen for an agent affecting nucleotide
 CC expression and gives easier monitoring of in vivo expression. The vector
 CC is used in the production of transgenic mammals for testing potential
 CC pharmaceutical or veterinary agents. pYMA4 is used to amplify YAC. The
 CC assay may be used to screen for agents useful in treatment of disturbance
 CC of circadian function, sleep disorders, eating disorders, premenstrual
 CC syndrome, autoimmune disorders, birth defects in women and/or sexual
 CC dysfunction. The agents thus detected may be used for treatment of
 CC disorders related to the expression pattern of a nucleotide such as those
 CC above. The vectors have more concentrated YAC DNA, which allows better
 CC and more reliable gene transfer. The presence of a reporter gene allows
 CC easy monitoring of in vivo expression and the vectors allow for gene
 CC overexpression (3-5 fold) and easy site determination. The pYMA4
 CC amplification vector does not contain the thymidine kinase gene, which
 CC causes male infertility in transgenic mice
 XX
 SQ Sequence 22 BP; 5 A; 11 C; 0 G; 6 T; 0 U; 0 Other;
 Query Match 69.1%; Score 15.2; DB 2; Length 22;
 Best Local Similarity 85.0%; Pred. No. 1.3e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 TCCCTCAGCTTCAGCCACAT 21
 DB 1 TTCTACCTCCTCAGCCACAT 20
 RESULT 5
 AAC85242
 ID AAC85242 standard; DNA; 22 BP.
 XX
 AC AAC85242;
 XX
 DT 22-MAR-2001 (first entry)
 XX
 DE Reverse primer D 26376 for determining size of YAC insert.
 XX
 KM Internal ribosomal entry site; IRS; yeast artificial chromosome; YAC;
 KM vector; centromere; telomere; origin of replication; transgenic;
 KM circadian function; sleep disorder; eating disorder;
 KM premenstrual syndrome; autoimmune disease; birth defect;
 KM sexual dysfunction; serotonin transporter; VIP2 receptor; SERT; VIPR2;
 KM polymerase chain reaction; PCR; primer; amplify; YAC 3508/D6;
 KM YAC HSC7526/V12; ss.
 XX
 OS Synthetic.

PN GB2350613-A.
 XX 06-DEC-2000.
 XX
 PF 17-AUG-2000; 2000GB-00020335.
 XX
 PR 28-NOV-1997; 97GB-00025311.
 PR 28-NOV-1997; 97GB-00025313.
 PR 20-MAR-1998; 98GB-00006072.
 PR 05-NOV-1998; 98GB-00024275.
 PR 27-NOV-1998; 98GB-00026126.
 XX
 PA (MED1-) MEDICAL RES COUNCIL.
 XX
 PI Shen S, Schedl A, Harmar AJ;
 XX
 DR WPI; 2001-034098/05.
 XX
 PT Transgenic organism for identifying potential therapeutic agents able to
 PT modulate gene expression, comprises a yeast artificial chromosome vector.
 XX
 PS Example; Page 54; 93pp; English.

CC The sequences given in AAC85227-50 are primers which were used to
 CC determine the size of the integrated YAC 35D8/D6 and YAC HSC7526/V12
 CC constructs in a transgenic founder animal. The constructs were prepared
 CC from novel yeast artificial chromosome (YAC) vectors each of which
 CC comprises a centromere, two telomeres, at least one origin of
 CC replication, an internal ribosomal entry site (IRES), and a selection
 CC gene that is specifically removable from the vector. The resulting YAC
 CC are used to produce transgenic organisms for use in screening for agents
 CC that can affect the expression pattern of a nucleotide sequence of
 CC interest (NOI) or the activity of its expression product. The identified
 CC agents are potentially useful as pharmaceutical and veterinary agents for
 CC treating disorders of circadian function; sleep or eating disorders;
 CC premenstrual syndrome; autoimmune diseases; birth defects in women and/or
 CC sexual dysfunction, and also as lead compounds for developing agents with
 CC other activities. YAC can also be used for expression, regulation and/or
 CC functional studies on NOI, in combination with other NOI, compounds or
 CC compositions. The new vectors provide high YAC copy numbers and allow
 CC easy monitoring (in vivo) of the expression pattern of NOI;
 CC (over)expression of NOI and a reporter gene; and determination of the
 CC sites where NOI is expressed. Incorporation of IRES allows expression of
 CC at least two nucleic acid sequences (e.g. NOI plus a reporter of more
 CC than one NOI)
 CC
 XX
 SQ Sequence 22 BP; 5 A; 11 C; 0 G; 6 T; 0 U; 0 Other;

Query Match 69.1%; Score 15.2; DB 5; Length 22;
 Best Local Similarity 85.0%; Pred. No. 1.3e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Gy 2 TCCCTCACCTCCAGCCACAT 21
 |||||
 Db 1 TTCAATCACTCCATCCACAT 20

RESULT 6

ACH62883
 ID ACH62883 standard; DNA, 25 BP.

ACH62883;
 AC

17-OCT-2003 (first entry)
 DT

DNA target sequence #12019 useful in array for genetic analyses.
 DE

Gene expression analysis; array; hybridisation; genetic variation;
 KW tag-labelled compound; gene family; in situ hybridisation;
 KW library screening; Southern hybridisation; northern hybridisation;
 KW dot-blot hybridisation; gene sequence; mutation detection;
 KW target sequence; probe; PCR; primer; ss.
 XX

OS Unidentified.
 XX
 PN US2003082596-A1.
 XX
 PD 01-MAY-2003.
 XX
 PF 08-AUG-2002; 2002US-00215112.
 XX
 PR 08-AUG-2001; 2001US-0311040P.
 XX
 PA (MITT/) MITTMANN M.
 XX
 PI Mittmann M;
 XX
 DR WPI; 2003-576608/54.
 XX
 PT New probe array useful e.g. for monitoring gene expression levels, for
 PT analyzing genetic variations, or for hybridizing tag-labeled compounds,
 PT comprises multiple nucleic acid probes.
 XX
 PS Claim 1; SEQ ID NO 12019; 9pp; English.

CC The present invention relates to nucleic acid sequences that are
 CC complementary to particular genes, and can be used as probes for a
 CC variety of analyses such as gene expression analysis. Each probe
 CC comprises 9 or more consecutive nucleotides from at least one of 14936
 CC nucleotide sequences defined in the patent, or their perfect sense match,
 CC sense mismatch, antisense match or antisense mismatch oligonucleotides.
 CC The probes may be used in an array comprising at least 10 distinct
 CC nucleic acid probes. The array is useful in monitoring gene expression
 CC levels by hybridisation to a DNA library, in analysing genetic
 CC variations, and in hybridising tag-labelled compounds. The probes are
 CC useful for identifying family members of a gene. The probes are also
 CC useful in in situ hybridisations, in screening cDNA or genomic libraries
 CC (or derived subclones) for additional clones containing segments of DNA
 CC that have been previously isolated and sequenced, in Southern, northern,
 CC or dot-blot hybridisation of genomic DNA to identify or detect the
 CC sequence of any gene or detect specific mutations in any gene, and in
 CC mapping the 5' termini of mRNA molecules by primer extensions. The
 CC nucleic acid sequences of the invention are also useful as PCR primers.
 CC The invention provides a large collection of nucleic acid sequences
 CC complementary to particular genes with a wide range of analytical uses.
 CC ACH50865-ACH65260 represent the target sequences of the invention. Note:
 CC The sequence data for this patent was obtained in electronic format
 CC directly from the USPTO web site at seqdata.uspto.gov/patidbidentry.htm
 CC
 XX

SQ Sequence 25 BP; 6 A; 11 C; 4 G; 4 T; 0 U; 0 Other;

Query Match 69.1%; Score 15.2; DB 9; Length 25;
 Best Local Similarity 85.0%; Pred. No. 1.3e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Gy 3 CCCTCACCTCCAGCCACATG 22
 |||||
 Db 1 CCTTCACTCCATCCACATG 20

RESULT 7

ACH63009
 ID ACH63009 standard; DNA, 25 BP.

ACH63009;
 AC

17-OCT-2003 (first entry)
 DT

DNA target sequence #12145 useful in array for genetic analyses.
 DE

Gene expression analysis; array; hybridisation; genetic variation;
 KW tag-labelled compound; gene family; in situ hybridisation;
 KW library screening; Southern hybridisation; northern hybridisation;
 KW dot-blot hybridisation; gene sequence; mutation detection;
 KW target sequence; probe; PCR; primer; ss.
 XX

OS Unidentified.
 XX US2003082596-A1.
 PN
 XX
 XX 01-MAY-2003.
 PD
 XX
 XX 08-AUG-2002; 2002US-00215112.
 PF
 XX
 XX 08-AUG-2001; 2001US-0311040P.
 PR
 XX
 XX (MITT/) MITTMANN M.
 PA
 XX
 XX Miltmann M;
 PI
 XX
 XX WPI; 2003-576608/54.
 DR
 XX
 XX
 XX New probe array useful e.g. for monitoring gene expression levels, for
 PT analyzing genetic variations, or for hybridizing tag-labeled compounds,
 PT comprising multiple nucleic acid probes.
 PT
 XX
 XX Claim 1; SEQ ID NO 12145; 9pp; English.
 PS
 XX
 XX The present invention relates to nucleic acid sequences that are
 CC complementary to particular genes, and can be used as probes for a
 CC variety of analyses such as gene expression analysis. Each probe
 CC comprises 9 or more consecutive nucleotides from at least one of 14936
 CC nucleotide sequences defined in the patent, or their perfect sense match,
 CC sense mismatch, antisense match or antisense mismatch oligonucleotides.
 CC The probes may be used in an array comprising at least 10 distinct
 CC nucleic acid probes. The array is useful in monitoring gene expression
 CC levels by hybridisation to a DNA library, in analysing genetic
 CC variations, and in hybridizing tag-labelled compounds. The probes are
 CC useful for identifying family members of a gene. The probes are also
 CC useful in *in situ* hybridisations, in screening cDNA or genomic libraries
 CC (or derived subclones) for additional clones containing segments of DNA
 CC that have been previously isolated and sequenced, in Southern, northern,
 CC or dot-blot hybridisation of genomic DNA to identify or detect the
 CC sequence of any gene or detect specific mutations in any gene, and in
 CC mapping the 5' termini of mRNA molecules by primer extensions. The
 CC nucleic acid sequences of the invention are also useful as PCR primers.
 CC The invention provides a large collection of nucleic acid sequences
 CC complementary to particular genes with a wide range of analytical uses.
 CC ACH50865-ACH55260 represent the target sequences of the invention. Note:
 CC The sequence data for this patent was obtained in electronic format
 CC directly from the USPTO web site at seqdata.uspto.gov/patid/patidntry.html
 CC
 XX
 XX Sequence 25 BP; 5 A; 11 C; 4 G; 5 T; 0 U; 0 Other;
 SQ
 Query Match 69.1%; Score 15.2; DB 9; Length 25;
 Best Local Similarity 85.0%; Pred. No. 1.3e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CCTTCACCTCCAGCCACATG 22
 |||||
 Db 1 CCTTCACCTCCATCCACATG 20
 |||||

RESULT 8
 AAV55821
 ID AAV55821 standard; DNA; 24 BP.
 XX
 AC AAV55821;
 XX
 DT 27-AUG-2003 (revised)
 DT 18-NOV-1998 (first entry)
 XX
 XX
 DE Multimerisation of minimal motifs using primer ZGY2.
 XX
 KM Fusion protein; stabilising polypeptide; proteolytic degradation;
 KM resistance; half-life; autoimmune disease; inflammation; *in vivo* imaging;
 KM IkappaB regulator protein; inflammatory bowel disease; *in vivo* imaging;
 KM nitroreductase protein; enzyme therapy; produg therapy; protease;
 KM cancer; pathological condition; minimal motif; PCR primer; ss.

XX
 OS Synthetic.
 OS Human herpesvirus 4.
 XX
 XX
 PN MO9822577-A1.
 XX
 XX 28-MAY-1998.
 PD
 XX
 XX 17-NOV-1997; 97WO-IB001508.
 PF
 XX
 XX 15-NOV-1996; 96US-0030986P.
 PR 25-JUN-1997; 97US-0048945P.
 XX
 XX (MASU/) MASUCCI M G.
 PA
 XX
 XX Masucci MG;
 PI
 XX
 XX WPI; 1998-312463/27.
 DR
 XX
 XX
 XX New fusion proteins resistant to proteolytic degradation - comprising a
 PT core protein with a stabilising polypeptide comprising a peptide sequence
 PT containing glycine repeats.
 PT
 XX
 XX Disclosure; Page 72; 120pp; English.
 PS
 XX
 XX Sequences shown in AAV55812 to AAV55827 represent primers used in the
 CC course of the invention for the multimerisation of minimal motifs. The
 CC invention provides a method for increasing the resistance of a core
 CC protein to proteolytic degradation that comprises linking or inserting
 CC onto or into the core protein a stabilising polypeptide of formula
 CC (G1ya)X(G1yb)Y(G1yc)Zln where G1ya, G1yb, Glyc are 1-6 sequential Gly
 CC residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met, Phe, Pro or Thr
 CC and n can be anything between 1-66. X, Y and Z need not be identical from
 CC n repeat to n repeat. Alternatively a nucleic acid encoding a stabilising
 CC polypeptide can be linked onto or inserted into a nucleic acid encoding a
 CC core protein. The fusion proteins of the invention are more resistant to
 CC degradation by proteases and, thus, have a longer half-life than the
 CC unfused core protein. The products can be used for treating autoimmune
 CC diseases, cancer and inflammation. In particular, the core protein may be
 CC an IkappaB regulator protein for the treatment of inflammatory bowel
 CC disease, or a nitroreductase protein which can activate nitro drugs in
 CC enzyme/prodrug therapy to treat cancer or other pathological conditions.
 CC The fusion proteins can also be used in diagnostic methods such as in
 CC *in vivo* imaging. (Updated on 27-AUG-2003 to correct OS field.)
 CC
 XX
 XX Sequence 24 BP; 5 A; 13 C; 2 G; 4 T; 0 U; 0 Other;
 SQ
 Query Match 66.4%; Score 14.6; DB 2; Length 24;
 Best Local Similarity 81.0%; Pred. No. 2.3e+04;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCCAGCCACAT 21
 |||||
 Db 4 CACCCGACCTCCAGCTCCAT 24
 |||||

RESULT 9
 ADY00944/c
 ID ADY00944 standard; DNA; 29 BP.
 XX
 AC ADY00944;
 XX
 DT 05-MAY-2005 (first entry)
 DT
 XX
 XX PCR primer 170 used to amplify human GALE region SNP DNA.
 DE
 XX
 XX SNP detection; breast tumor; endocrine disease;
 KM gynecology and obstetrics; neoplasia; cytostatic; metastasis;
 KM gene therapy; RNA interference; UDP-galactose 4-epimerase; GALE;
 KM UDP-glucose 4-epimerase; lysophospholipase; HT014; LOC148902; LYPLA2; ss;
 KM PCR primer.
 XX
 OS Homo sapiens.

CC genes designed from repetitive peptide sequences, such as glycoproteins
CC (including the peptide sequences of gum arabic glycoprotein (GAGP)).
CC ABL51730 to ABL51849 and ABB78401 to ABB78544 represent sequences used in
CC the exemplification of the present invention

SO Sequence 30 BP; 6 A, 19 C, 0 G, 5 T, 0 U, 0 Other;

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 2.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCACAT 21
DB 5 CACCTTCAGCTCCAGCCCAT 25

RESULT 12

ABL51739
ID ABL51739 standard; DNA; 30 BP.

AC ABL51739;

DT 09-JUL-2002 (first entry)

XX Hydroxyproline-rich glycoprotein (HRGP) related linker SEQ ID NO:38.

XX Plant; Gum arabic glycoprotein; GAGP; hydroxyproline-rich glycoprotein;
XX HRGP; repetitive proline-rich protein; RRP; arabinogalactan protein;
XX AGP; plant gum; PCR primer; linker; ss.

OS Acacia senegal.
XX Synthetic.

XX WO200178503-A2.

XX 25-OCT-2001.

XX 12-APR-2001; 2001WO-US012336.

XX 12-APR-2000; 2000US-00547693.

XX (UYOH-) UNIV OHIO.

XX Kieleszewski MJ;

XX WPI; 2002-041307/05.

XX Nucleic acids and proteins useful for producing hydroxy-proline rich
XX glycoproteins in plants.

XX Example 2; Page 53; 326pp; English.

XX The present invention describes synthetic genes encoding plant gums and
XX other hydroxyproline (Hyp)-rich glycoproteins (HRGP) and the nucleic
XX acids that encode them. The nucleic acids, proteins and methods from the
XX present invention may be used to produce HRGP, repetitive proline-rich
XX proteins (RRP) and arabinogalactan-proteins (AGP) in plants via
XX recombinant methodologies. Also described is the expression of synthetic
XX genes designed from repetitive peptide sequences, such as glycoproteins
XX (including the peptide sequences of gum arabic glycoprotein (GAGP)).
XX ABL51730 to ABL51849 and ABB78401 to ABB78544 represent sequences used in
XX the exemplification of the present invention

SO Sequence 30 BP; 6 A, 19 C, 0 G, 5 T, 0 U, 0 Other;

Query Match 66.4%; Score 14.6; DB 6; Length 30;
Best Local Similarity 81.0%; Pred. No. 2.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCACAT 21
DB 5 CACCTTCAGCTCCAGCCCAT 25

RESULT 13

ADU09143
ID ADU09143 standard; DNA; 30 BP.

AC ADU09143;

DT 27-JAN-2005 (first entry)

XX Ser-Hyp4 gene cassette oligonucleotide #7.

XX Gum arabic glycoprotein; GAGP; Hydroxyproline-rich glycoprotein; HRGP;
XX repetitive proline-rich protein; RRP; arabinogalactan-protein; AGP;
XX plant gum; elastin; extensin; thickener; emulsifier; stabiliser;
XX surfactant; flavour fixative; food industry; cosmetic; ss.

OS Synthetic.

XX WO2004094590-A2.

XX 04-NOV-2004.

XX 13-APR-2004; 2004WO-US011174.

XX 16-APR-2003; 2003US-00418032.

XX (UYOH-) UNIV OHIO.

XX Kieleszewski MJ;

XX WPI; 2004-784900/77.

XX Plant gums (especially gum arabic) useful in the food and cosmetics
XX industries are produced by expression of synthetic genes designed from
XX the repetitive peptide sequences of plant gum glycoproteins.

XX Example 2; SEQ ID NO 38; 179pp; English.

XX The invention relates to an isolated polypeptide (I) comprising
XX alternating rigid and non-rigid modules, where the rigid modules are
XX directed by contiguous hydroxyproline residues and the non-rigid modules
XX are directed by non-contiguous hydroxyproline residues. The polypeptides
XX are synthetic gums based on motifs and repeats found in Gum arabic
XX glycoprotein (GAGP), Hydroxyproline-rich glycoproteins (HRGP), repetitive
XX proline-rich proteins (RRP) and arabinogalactan-proteins (AGP). Also
XX included are a polypeptide (II) (comprising a first elastin module
XX flanked on the N-terminal side by a first extensin module having a first
XX crosslinkage motif, and a second elastin module flanked on the C-terminal
XX side by a second extensin module having a second crosslinkage motif) and
XX a nucleic acid encoding a polypeptide variant of Lycopodium esculentum
XX arabinogalactan protein (Le-AGP-1). In polypeptide (I), the rigid module
XX comprises at least two contiguous hydroxyproline residues and is
XX arabinosylated. Polypeptide (II) further comprises a glycoprotein having a
XX crosslinkage motif. The crosslinkage motif comprises tyrosine, preferably
XX the crosslinkage motif is Val-Tyr-Lys. The tyrosine forms an
XX intramolecular or intermolecular crosslink. In polypeptide (II), the
XX first elastin module is repeated six times and the second elastin module
XX is repeated three times. The first and/or second crosslinkage motifs
XX preferably comprise Val-Tyr-Lys and contiguous hydroxyproline residues.
XX (II) further comprises a central stretch of rigid arabinosylated Ser-O-
XX O-O (where O is hydroxyproline, appearing as ADU09143) repeats flanked on
XX either side by the first and the second elastin modules. The Le-AGP-1
XX variant encoded by lacks a glycosylphosphatidylinositol anchor signal
XX sequence and an internal lysine rich region. The variant encoded by the
XX nucleic acid further comprises an enhanced green fluorescent protein
XX fusion glycoprotein. The nucleic acid is expressed in Nicotiana tabacum.
XX The polypeptide (I) is useful as a thickener, emulsifier, stabiliser,
XX surfactant, or flavour fixative in the food industry. It is also useful
XX in cosmetics. The polypeptide (I) has an anomalously low viscosity. The
XX present sequence encodes a cassette peptide comprising one or more
XX consensus motifs or repeat units derived from GAGP, AGP, RRP or HRGP
XX used to make a polypeptide of the invention.

SQL Sequence 30 BP; 6 A; 19 C; 0 G; 5 T; 0 U; 0 Other;

Query Match 66.4%; Score 14.6; DB 13; Length 30;
Best Local Similarity 81.0%; Pred. No. 2.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCTCCACCAT 21
DB 5 CACCTTCACCTCCACCCCAT 25

RESULT 14
ADU09116
ID ADU09116 standard; DNA; 30 BP.

AC ADU09116;
DT 27-JAN-2005 (first entry)

DE Hydroxyproline-rich glycoprotein, HPGP, conserved motif DNA.

KM Gum arabic glycoprotein; GAGP; Hydroxyproline-rich glycoprotein; HRGP;
KM repetitive proline-rich protein; RRP; arabinogalactan-protein; AGP;
KM plant gum; elastin; extensin; thickener; emulsifier; stabiliser;
KM surfactant; flavour fixative; food industry; cosmetic; ss.

OS Unidentified.
XX
XX WO2004094590-A2.
XX
XX 04-NOV-2004.
XX
XX 13-APR-2004; 2004WO-US011174.
XX
XX PF
XX 16-APR-2003; 2003US-00418032.
XX
XX PR
XX (UYOH-) UNIV OHIO.
XX
XX PA
XX Kieliszewski MJ;
XX
XX WPI; 2004-784900/77.
XX
XX DR
XX
XX PT Plant gums (especially gum arabic) useful in the food and cosmetics
XX industries are produced by expression of synthetic genes designed from
XX the repetitive peptide sequences of plant gum glycoproteins.
XX
XX PS Disclosure; SEQ ID NO 10; 179pp; English.

XX The invention relates to an isolated polypeptide (I) comprising
XX alternating rigid and non-rigid modules, where the rigid modules are
XX directed by contiguous hydroxyproline residues and the non-rigid modules
XX are directed by non-contiguous hydroxyproline residues. The polypeptides
XX are synthetic gums based on motifs and repeats found in gum arabic
XX glycoprotein (GAGP), Hydroxyproline-rich glycoproteins (HRGP), repetitive
XX proline-rich proteins (RRP) and arabinogalactan-proteins (AGP). Also
XX included are a polypeptide (II) (comprising a first elastin module
XX flanked on the N-terminal side by a first extensin module having a first
XX crosslinkage motif, and a second elastin module flanked on the C-terminal
XX side by a second extensin module having a second crosslinkage motif) and
XX a nucleic acid encoding a polypeptide variant of Lycopodium seculentum
XX arabinogalactan protein (Le-AGP-1). In polypeptide (I), the rigid module
XX comprises at least two contiguous hydroxyproline residues and is
XX arabinosylated. Polypeptide (I) further comprises a glycoprotein having a
XX crosslinkage motif. The crosslinkage motif comprises tyrosine, preferably
XX the crosslinkage motif is Val-Tyr-Lys. The tyrosine forms an
XX intramolecular or intermolecular crosslink. In polypeptide (II), the
XX first elastin module is repeated six times and the second elastin module
XX is repeated three times. The first and/or second crosslinkage motif
XX preferably comprises Val-Tyr-Lys and contiguous hydroxyproline residues.
XX (II) further comprises a central stretch of rigid arabinosylated Ser-O-O-
XX (where O is Hydroxyproline, appearing as ADU09308) repeats flanked on
XX either side by the first and the second elastin modules. The Le-AGP-1
XX variant encoded by lacks a glycosylphosphatidylinositol anchor signal

CC sequence and an internal lysine rich region. The variant encoded by the
CC nucleic acid further comprises an enhanced green fluorescent protein
CC fusion glycoprotein. The nucleic acid is expressed in Nicotiana tabacum.
CC The polypeptide (I) is useful as a thickener, emulsifier, stabiliser,
CC surfactant, or flavour fixative in the food industry. It is also useful
CC in cosmetics. The polypeptide (I) has an anomalously low viscosity. The
CC present sequence encodes a consensus motif or repeat unit derived from
CC GAGP, AGP, RRP or HRGP.
XX
XX SQL Sequence 30 BP; 6 A; 19 C; 0 G; 5 T; 0 U; 0 Other;

Query Match 66.4%; Score 14.6; DB 13; Length 30;
Best Local Similarity 81.0%; Pred. No. 2.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCTCCACCAT 21
DB 5 CACCTTCACCTCCACCCCAT 25

RESULT 15
AAL49623/C
ID AAL49623 standard; DNA; 21 BP.

AC AAL49623;
DT 27-NOV-2002 (first entry)

DE Tumour differentiation effecting protein TL4 related PCR primer #26.

KM Mouse; tumour differentiation; rhabdomyosarcoma; leiomyosarcoma; rat; ss;
KM muscular dystrophy; uterine myoma; cytostatic; plasmic change; TL4;
KM human; PCR; primer.

OS Unidentified.
XX
XX WO20026049-A1.
XX
XX PN
XX 29-AUG-2002.
XX
XX PD
XX 21-FEB-2002; 2002WO-JP001536.
XX
XX PF
XX 23-FEB-2001; 2001JP-00049450.
XX
XX PR
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX PA
XX Hikichi Y, Shintani Y, Matsui H;
XX
XX PI
XX WPI; 2002-674894/72.
XX
XX DR
XX
XX PT Plasmic change agents and antibodies to them for diagnosis and treatment
XX of tumors of muscle tissue and of muscular dystrophy.
XX
XX PS Example 7; Page 129; 136pp; Japanese.

XX The present invention relates to plasmic change agents with cell
XX differentiation activity containing protein TL4. These can be used in the
XX treatment, prevention and diagnosis of rhabdomyosarcoma, leiomyosarcoma,
XX muscular dystrophy and uterine myeloma. The present sequence is a PCR
XX primer used in the exemplification of the invention
XX
XX SQL Sequence 21 BP; 5 A; 2 C; 11 G; 3 T; 0 U; 0 Other;

Query Match 65.5%; Score 14.4; DB 6; Length 21;
Best Local Similarity 93.8%; Pred. No. 2.7e+04;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCACCTCCACG 16
DB 18 CTCCTCACCTCCATC 3

Search completed: March 25, 2006, 19:02:17

Mon Mar 27 08:53:26 2006

us-10-798-652-6.sz1m30.png

Page 9

Job time : 375.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 25, 2006, 13:43:43 ; Search time 3041.5 Seconds
(without alignments)
338.424 Million cell updates/sec

Title: US-10-798-652-6
Perfect score: 22
Sequence: 1 ctccctcacctccagcacatg 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 52094

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: gb_eac1:*
2: gb_eac2:*
3: gb_eac3:*
4: gb_hrc:*
5: gb_eac4:*
6: gb_eac5:*
7: gb_eac6:*
8: gb_eac7:*
9: gb_ga81:*
10: gb_ga82:*
11: gb_ga83:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14	63.6	29	10	CL676355 PR10118b_
C 2	13.6	61.8	25	8	D21049 HMGSO2034
C 3	13.2	60.0	22	9	AZ980990 2M0258F17
C 4	13	59.1	24	9	AZ655652 1M0530D13
C 5	13	59.1	28	1	AA864650 ob37b09.s
C 6	12.8	58.2	29	9	AZ498825 1M0336J14
C 7	12.6	57.3	22	9	AZ778032 2M0013A11
C 8	12.6	57.3	23	9	AZ785027 2M0028H03
C 9	12.6	57.3	25	10	CG730897
C 10	12.2	55.5	19	9	AZ381798 1M0138G01
C 11	12.2	55.5	19	9	AZ782026 2M0021I23
C 12	12.2	55.5	27	9	AZ501659 1M0340H12
C 13	12	54.5	23	9	AZ599021 1M0414F08
C 14	12	54.5	24	9	AZ762096 1M0556J18
C 15	12	54.5	25	1	AI416870 sal1b03.x
C 16	12	54.5	25	9	AZ462642 1M0269M09
C 17	12	54.5	29	9	AZ332711 1M0061N14
C 18	12	54.5	29	9	AZ788256 2M0035A17
C 19	12	54.5	29	10	CG477475
C 20	11.8	53.6	21	9	AZ348213 1M0084N15
C 21	11.8	53.6	21	9	AZ592392 1M0403J14
C 22	11.6	52.7	19	9	AZ453412 1M0254I10

23	11.6	52.7	22	8	D21051	D21051 HMGSO2036
C 24	11.6	52.7	23	9	AZ331349	AZ331349 1M0059F17
C 25	11.6	52.7	23	9	AZ410218	AZ410218 1M0182F23
C 26	11.6	52.7	27	9	AZ623416	AZ623416 1M0461C08
C 27	11.6	52.7	29	9	AZ580965	AZ580965 1M0369C15
C 28	11.6	52.7	30	9	AZ833458	AZ833458 2M0115L02
C 29	11.4	51.8	24	10	AG202658	AG202658 Pan t10g1
C 30	11.4	51.8	25	9	AZ471181	AZ471181 1M0285H11
C 31	11.2	50.9	23	9	AZ374746	AZ374746 1M0127D15
C 32	11.2	50.9	29	9	AZ812552	AZ812552 2M0079K13
C 33	11	50.0	19	1	A1538541	A1538541 tcd08a11.x
C 34	11	50.0	22	2	BG926061	BG926061 HNC23-1-E
C 35	11	50.0	23	9	AZ654389	AZ654389 1M0528H06
C 36	11	50.0	24	9	AZ404465	AZ404465 1M0172P09
C 37	11	50.0	25	1	A1620280	A1620280 tu93e03.x
C 38	11	50.0	25	6	CP298174	CP298174 7LEAF--01
C 39	11	50.0	25	9	AZ348233	AZ348233 1M0084G04
C 40	11	50.0	26	10	CL650847	CL650847 PR10110C
C 41	11	50.0	27	9	AZ439789	AZ439789 1M0230D04
C 42	11	50.0	28	8	N81476	N81476 TGBST-y47h0
C 43	11	50.0	28	9	AZ937148	AZ937148 2M0195A18
C 44	11	50.0	29	9	AZ666737	AZ666737 1M0549O06
C 45	11	50.0	29	9	AZ780387	AZ780387 2M0017N12

ALIGNMENTS

RESULT 1
LOCUS CL676355/29 bp DNA linear GSS 09-JUL-2004
DEFINITION PR10118b E09 2 - PR10118b.BR (29) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.

ACCESSION CL676355
VERSION CL676355.1 GI:50181863
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasterida; Pristionchus.

AUTHORS Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE 1 (bases 1 to 29)
JOURNAL ApadB: an Acids database for the nematode satellite organism Pristionchus pacificus
PUBLISHED Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT 14681447

CONTACT: Sommer RJ

Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498

Email: ralf.sommer@cuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
Seq primer: 17
Class: fosmid ends.

FEATURES
source location/Qualifiers

1..29
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus var. California"
/note="Vector: pBg1fos-5 Fosmid vector"

ORIGIN
Query Match 63.6%; Score 14; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.4e+05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	4	CCTCACCCTCCAGCC	17
Db	15	CCTCACCTTCACGCC	2
RESULT 2			
D21049			25 bp mRNA linear EST 30-JUL-1996
LOCUS		HMGSSD2034 Human promyelocyte Homo sapiens cDNA clone mp038 3'	
DEFINITION		mRNA sequence.	
ACCESSION		D21049	
VERSION		D21049.1	GI:504869
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
TITLE		Homidae; Homo.	
JOURNAL		1 (bases 1 to 25)	
COMMENT		Okubo, K., Fukushima, A., Yoshii, J., Niijima, T., Kojima, Y., Yoshihara, H., Arimoto, J. and Matsubara, K. Gene expression of human promyelocytic cell line HL60 before and after induction of differentiation. A new application of 3'directed cDNA sequencing Unpublished (1993) Contact: Okubo, K., Fukushima, A., Yoshii, J., Niijima, T., Kojima, Y., Yoshihara, H., Arimoto, J. and Matsubara, K. Institute for Molecular and Cellular Biology Osaka University 3-1 Yamada-oka, Suita, Osaka 565, Japan.	
FEATURES		Location/Qualifiers	
SOURCE		1..25	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="mp0338"	
		/clone_lib="Human promyelocyte"	
		/note="Female, adult, cell_line = HL60, cell_type = promyelocyte."	
ORIGIN			
Query Match		61.8%; Score 13.6; DB 8;	
Best Local Similarity		80.0%; Pred. No. 3.5e+05;	
Matches	16;	Conservative	0; Mismatches 4; Indels 0; Gaps 0;
OY	1	CTCCTCACCCTCCAGCCACA	20
Db	4	CGCCCCCACCCTCCACCTCA	23
RESULT 3			
AZ980990/c		22 bp DNA linear GSS 27-APR-2001	
LOCUS		ZM0258F17F Mouse 10kb plasmid UUCG2M library Mus musculus genomic	
DEFINITION		clone UUGCGZM0258F17 F, genomic survey sequence.	
ACCESSION		AZ980990	
VERSION		AZ980990.1	GI:13852217
KEYWORDS		GSS.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;	
		Sciurognathi; Muridae; Murinae; Murinae; Murinae.	
		1 (bases 1 to 22)	
		Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Tetam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Niederhausen, A., Rose, R., Stokes, R., Tingey, A., von Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts Unpublished (2000) Contact: Robert B. Weiss University of Utah Genome Center	

University of Utah
 Rm. 308, Biomedical Polymer Research Bldg., 20 S. 2030 E., SIC, UT
 84112, USA
 Tel.: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0258 row: F column: 17
 Seq primer: CATTGTAAACGACGCGCACT
 Class: plasmid ends
 High quality sequence stop: 22.
 Location/Qualifiers
 1. .22
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0258F17"
 /sex="Female"
 /lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
 /clone_1lb="Mouse 10kb plasmid UUGC2M library"
 /note="Vector: FMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMW2 (g1473214|g14739072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

Query Match 60.0%; Score 13.2; DB 9; Length 22;
 Best Local Similarity 83.3%; Pred. No. 5.2e+05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CCTCTACCTCCAGCCACA 20
 ||| ||||| |||||
 Db 22 CCCCACCTCCACACACA 5

RESULT 4
 A2655652 24 bp DNA linear GSS 14-DEC-2000
 LOCUS IM0530D13R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
 DEFINITION clone UUGC2M0530D13 R, genomic survey sequence.
 ACCESSION A2655652
 VERSION A2655652.1 GI:11792798
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 24)
 Dunn, D., Aoyagi, A., Barber, M., Beacom, T., Duval, B., Hamil, C.,
 Irlam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausen, A. and Wright, D., Weis, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished. (2000)
 Contact: Robert B. Weis

JOURNAL
 COMMENT

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunne@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0530 row: D column: 13

Seq primer: CACACAGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 24.

FEATURES

source

Location/Qualifiers

1..24

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="UUCG1M0530D13"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUCG1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pMD2 (g14732114[gblAF129072.1]), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

ORIGIN

Query Match 59.1%; Score 13; DB 9; Length 24;

Best Local Similarity 76.2%; Pred. No. 6.3e+05;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TCCCTCACTCCAGCCACATG 22

Db 23 TACCTGACATCCACCCACAG 3

RESULT 5

AA864650/c

LOCUS

DEFINITION

AA864650 28 bp mRNA linear EST 13-MAY-1998

0317009.g1 NCI CGAP Kid6 Homo sapiens cDNA IMAGE:1459961.3'

similar to SW:ATPO_EOVIN P13620 ATP SYNTHASE D CHAIN, MITOCHONDRIAL

;; mRNA sequence.

AA864650

AA864650.1 GI:2958963

EST.

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 28)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

JOURNAL

COMMENT

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Stratagene, Inc.

CDNA Library Arrayed by: Greg Lemon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/INL at:

www.bio.linnl.gov/bdip/image/image.html

Insert Length: 952 Std Error: 0.00

Seq primer: -40ml3 fwd. RT from Amersham

High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers

1..28

/organism="Homo sapiens"

/mol_type="RNA"

/db_xref="taxon:9606"

/clone="IMAGE:1459961"

/sex="mixed"

/issue_type="Kidney tumor"

/lab_host="SOLR (kanamycin resistant)"

/clone_lib="NCI CGAP Kid6"

/note="Organ: kidney; Vector: Bluescript SK-; Site 1:

ECORI; Site 2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'

GAATTCGACACAG 3' 3' adaptor sequence: 5'

CTGACGATTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."

ORIGIN

Query Match 59.1%; Score 13; DB 1; Length 28;

Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCCCTCACTCCAG 15

Db 27 CCCCTCACTCCAG 15

RESULT 6

AZ498825/c

LOCUS

DEFINITION

AZ498825 29 bp DNA linear GSS 05-OCT-2000

IM0336J1F Mouse 10kb plasmid UUCG1M library Mus musculus genomic

clone UUCG1M0336J1 F, genomic survey sequence.

AZ498825

AZ498825.1 GI:10677040

GSS.

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 29)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,

Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von

Niederhausen,A. and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunne@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0336 row: J column: 14

Seq primer: CGTTGTAACGACGCGCACT

Class: plasmid ends

High quality sequence stop: 29.

Location/Qualifiers

1..29

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCGM0336J14"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_1ib="Mouse 10kb plasmid UUGCM library"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 58.2%; Score 12.8; DB 9; Length 29;
Best Local Similarity 87.5%; Pred. No. 7.7e+05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCCTACCTCCAGCC 17
|||||
Db 24 TCCCTACCTCCAGCC 9

RESULT 7
A2778032/c 22 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0013A11F Mouse 10kb plasmid UUGCM library Mus musculus genomic
DEFINITION clone UUGCM0013A11 F, genomic survey sequence.
ACCESSION A2778032
VERSION A2778032.1 GI:12907248
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 22)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Petersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

TITLE Unpublished (2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0013 row: A column: 11
Seq primer: CGTTGTAACAGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 22.
Location/Qualifiers

FEATURES

source

1. 22
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCGM0013A11"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_1ib="Mouse 10kb plasmid UUGCM library"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 57.3%; Score 12.6; DB 9; Length 22;
Best Local Similarity 78.9%; Pred. No. 9.2e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTACCTCCAGCCAC 19
|||||
Db 22 CCCCTACCTCCAGCCAC 4

RESULT 8
A2785027/c 23 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0028H03R Mouse 10kb plasmid UUGCM library Mus musculus genomic
DEFINITION clone UUGCM0028H03 R, genomic survey sequence.
ACCESSION A2785027
VERSION A2785027.1 GI:12921357
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Petersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

TITLE Unpublished (2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0028 row: H column: 03
Seq primer: CACACAGAAACAGCTATGAC
Class: plasmid ends
High quality sequence stop: 23.

ORIGIN

Qy 3 CCTCACCCTCCAGCCAC 1
||| ||| ||| |||
Db 19 CCCCACACCACCGCAC 3

ACCESSION	AZ782026	GI:12915307
VERSION	AZ782026.1	
KEYWORDS	GSS.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Relly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

TITLE Niederhausem, A. and Wright, D., Weiss, R. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

University of Utah
University of Utah
Rm. 308, Biomedical Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0021 row: 1 column: 23
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.

1. .19

```

"organism="Mus musculus"
"mol_type="genomic DNA"
"strain="CS7BL/6J"
"db_xref="taxon:10090"
"clone="U06C2M0021123"
"sex="Male"
"lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
"clone_lib="mouse 10kb plasmid U06C2M library"

```

ORIGIN

Qy	1	CTCCCTCACCCTCGAGCC	17
Db	17	CACCCGACCTCTGCCC	1

ACCESSION	A2501699	
VERSION	A2501699.1	GI:10683015
KEYWORDS	GSS.	
SOURCE	Mus musculus	(house mouse)
ORGANISM	Mus musculus	

Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLCC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0340 row: H column: 12
Seq primer: CACACAGGAACACGCTTGACC
Class: plasmid ends
High quality sequence stop: 27.

Source	1. .27
--------	--------

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MUSC1M0340H12"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

```

/clone_11b="Mouse 10kb plasmid UGCG1M library"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 55.5%; Score 12.2; DB 9; Length 27;
Best Local Similarity 82.4%; Pred. No. 1.4e+06;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 4 CCTCAGCTCCAGCCACA 20
Db 25 CCTCAGCTCCAGCCACA 9

RESULT 13

AZ599021/c 23 bp DNA linear GSS 13-DEC-2000
LOCUS 1M0414F08F Mouse 10kb plasmid UGCG1M library Mus musculus genomic
DEFINITION clone UGCG1M0414F08 F, genomic survey sequence.
ACCESSION AZ599021
VERSION AZ599021.1 GI:11721211
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, B., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std. Error: 0.00
Plate: 0414 row: F column: 08
Seq primer: CGTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 23.
Location/Qualifiers

FEATURES

1..23
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG1M0414F08"
/sex="Male"

ORIGIN

Query Match 54.5%; Score 12; DB 9; Length 23;
Best Local Similarity 75.0%; Pred. No. 1.6e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CTCCTCAGCTCCAGCCACA 20
Db 21 CTACCCAGCTCCAGCCACA 2

RESULT 14

AZ762096/c 24 bp DNA linear GSS 16-FEB-2001
LOCUS AZ762096
DEFINITION 1M0556J18 Mouse 10kb plasmid UGCG1M library Mus musculus genomic
clone UGCG1M0556J18 R, genomic survey sequence.
ACCESSION AZ762096
VERSION AZ762096.1 GI:12871740
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 24)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, B., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std. Error: 0.00
Plate: 0556 row: J column: 18
Seq primer: CACACAGCAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 24.
Location/Qualifiers

FEATURES

1..24
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG1M0556J18"

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCM library"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male); was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid p1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

ORIGIN

```

Query Match          54.5%; Score 12; DB 9; Length 24;
Best Local Similarity 75.0%; Pred. No. 1.6e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      1 CTGCGTCACCTCGAGCCACA 20
          ||||| ||||| |||||
DB      24 CCCCCCCCCGCGCAGCCACA 5

```

RESULT 15

```

LOCUS      A1416870                25 bp      mRNA      linear      EST 23-UTL-2004
DEFINITION ba19603.x1 Gm-c1005 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1005-102 3' similar to TR:Q39599 Q39599 EXTENSIN.1; mRNA
            sequence.

```

ACCESSION

A1416870

VERSION

A1416870.1 GI:4260374

KEYWORDS

EST.

SOURCE

Glycine max (soybean)

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

1 (bases 1 to 25)

REFERENCE

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,
Bowers, J., Pearson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished (1999)

COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. GENOME SYSTEMS

CLONE ID: Gm-c1005-102 Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand this clone is

available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163, email: info@biogeneticservices.com)

Seq primer: T7 ET from Amersham

High quality sequence stop: 1
POLYA=No.

FEATURES

source

```

Location/Qualifiers
1..25
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1005-102"
/lab_host="XL10-Gold"
/clone_lib="Gm-c1005"
/note="Vector: pBluescript II XE; Site 1: EcoRI; Site 2:
XhoI. This library was constructed by Dr. Randy C.
Shoemaker and Dr. John Erpelding, USDA-ARS Agronomy
Department, G401 Agronomy Hall, Iowa State University,
 Ames, IA 50011-1010 Phone: 515-294-6233 FAX: 515-294-2299
email: rcshoe@iastate.edu"

```

ORIGIN

```

Query Match          54.5%; Score 12; DB 1; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.7e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      1 CTGCGTCACCTCGAGCCACA 20
          ||||| ||||| |||||
DB      6 CCCCCCCCCCTCACCCCA 25

```

Search completed: March 25, 2006, 15:28:17
Job time : 3045.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 12:58:33 ; Search time 129.5 Seconds
(without alignments)
301.980 Million cell updates/sec

Title: US-10-798-652-6

Perfect score: 22
Sequence: 1 ctccctcaccctccagccacatg 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 1026780

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents, NA:*

- 1: /cgn2_6/pdata/1/ina/1_COMB.seq:*
- 2: /cgn2_6/pdata/1/ina/5_COMB.seq:*
- 3: /cgn2_6/pdata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/pdata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/pdata/1/ina/H_COMB.seq:*
- 6: /cgn2_6/pdata/1/ina/PC/US_COMB.seq:*
- 7: /cgn2_6/pdata/1/ina/PP_COMB.seq:*
- 8: /cgn2_6/pdata/1/ina/RG_COMB.seq:*
- 9: /cgn2_6/pdata/1/ina/Backfillseq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.6	66.4	24	2	US-08-529-190B-16
2	14.6	66.4	30	3	US-09-119-507B-10
3	14.6	66.4	30	3	US-09-119-507B-38
4	14.6	66.4	30	3	US-08-897-556A-10
5	14.6	66.4	30	3	US-08-897-556A-38
6	14.6	66.4	30	3	US-09-547-693-10
7	14.6	66.4	30	3	US-09-547-693-38
8	14.2	64.5	24	2	US-08-529-190B-14
9	14.2	64.5	24	3	US-07-792-600-8
10	14.2	64.5	24	3	US-09-157-021-8
11	14.2	64.5	24	3	US-09-156-842-8
12	14.2	64.5	24	3	US-09-591-514-8
13	14.2	64.5	24	3	US-09-887-145-6
14	13.8	63.7	20	3	US-09-674-824-5
15	13.6	61.8	20	3	US-09-487-368A-216
16	13.6	61.8	20	3	US-09-658-679A-52
17	13.6	61.8	20	3	US-09-629-644A-216
18	13.6	61.8	20	3	US-09-629-644A-216
19	13.6	61.8	24	2	US-08-529-190B-4
20	13.6	61.8	24	2	US-08-529-190B-7
21	13.6	61.8	25	3	US-09-396-196G-40878
22	13.6	61.8	25	3	US-09-396-196G-101880
23	13.6	61.8	27	2	US-08-460-890A-19
24	13.6	61.8	27	2	US-08-460-890A-22

C 25	13.6	61.8	27	3	US-08-167-641C-19	Sequence 19, Appl
C 26	13.6	61.8	27	3	US-08-167-641C-22	Sequence 22, Appl
C 27	13.6	61.8	27	3	US-08-460-971A-19	Sequence 19, Appl
C 28	13.6	61.8	27	3	US-08-460-971A-22	Sequence 22, Appl
C 29	13.6	61.8	27	3	US-08-462-040-19	Sequence 19, Appl
C 30	13.6	61.8	27	3	US-08-462-040-22	Sequence 22, Appl
C 31	13.4	60.9	25	3	US-09-396-196G-59434	Sequence 59434, A
C 32	13.2	60.0	19	3	US-09-358-972-221	Sequence 221, App
C 33	13.2	60.0	19	3	US-09-406-064-28	Sequence 28, Appl
C 34	13.2	60.0	19	3	US-09-430-615-58	Sequence 58, Appl
C 35	13.2	60.0	19	3	US-09-406-065-62	Sequence 62, Appl
C 36	13.2	60.0	19	3	US-09-383-316-78	Sequence 78, Appl
C 37	13.2	60.0	19	3	US-09-788-847-281	Sequence 281, App
C 38	13.2	60.0	19	3	US-09-790-417-221	Sequence 221, App
C 39	13.2	60.0	22	9	5219727-31	Patent No. 5219727
C 40	13.2	60.0	24	2	US-08-529-190B-2	Sequence 2, Appl
C 41	13.2	60.0	24	2	US-08-529-190B-5	Sequence 5, Appl
C 42	13.2	60.0	24	2	US-08-463-978-2	Sequence 2, Appl
C 43	13.2	60.0	24	2	US-08-463-978-5	Sequence 5, Appl
C 44	13.2	60.0	24	6	PCT-US94-05467-2	Sequence 2, Appl
C 45	13.2	60.0	24	6	PCT-US94-05467-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-529-190B-16
Sequence 16, Application US/08529190B
Patent No. 5833991
GENERAL INFORMATION:
APPLICANT: Masucci, Maria G.
TITLE OF INVENTION: GLYCINE-CONTAINING SEQUENCES
TITLE OF INVENTION: CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: One Financial Center
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/529,190B
FILING DATE: 15-SEP-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE9501324-9
FILING DATE: 10-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/522,595
FILING DATE: 01-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen A
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/53015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
US-08-529-190B-16

Query Match 66.4%; Score 14.6; DB 2; Length 24;
Best Local Similarity 81.0%; Pred. No. 3.8e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
DB 4 CACCCGACCTCCAGCTCCAT 24

RESULT 2
US-09-119-507B-10
; Sequence 10, Application US/09119507B
; Patent No. 6548642
; GENERAL INFORMATION:
; APPLICANT: Kieliszewski, Marcia J.
; TITLE OF INVENTION: No. 6548642el Synthetic Genes for Plant Gums
; FILE REFERENCE: OHU-03417
; CURRENT APPLICATION NUMBER: US/09/119,507B
; CURRENT FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-119-507B-10

Query Match 66.4%; Score 14.6; DB 3; Length 30;
Best Local Similarity 81.0%; Pred. No. 3.9e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
DB 5 CACCTTCACTCCACCCCAT 25

RESULT 3
US-09-119-507B-38
; Sequence 38, Application US/09119507B
; Patent No. 6548642
; GENERAL INFORMATION:
; APPLICANT: Kieliszewski, Marcia J.
; TITLE OF INVENTION: No. 6548642el Synthetic Genes for Plant Gums
; FILE REFERENCE: OHU-03417
; CURRENT APPLICATION NUMBER: US/09/119,507B
; CURRENT FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-119-507B-38

Query Match 66.4%; Score 14.6; DB 3; Length 30;
Best Local Similarity 81.0%; Pred. No. 3.9e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
DB 5 CACCTTCACTCCACCCCAT 25

RESULT 4
US-08-897-556A-10
; Sequence 10, Application US/08897556A
; Patent No. 6570062
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA J.

; TITLE OF INVENTION: SYNTHETIC GENES FOR PLANT GUMS AND OTHER
; TITLE OF INVENTION: HYDROXYPROLINE-RICH GLYCOPROTEINS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,556A
; FILING DATE: 21-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OHU-02908
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
US-08-897-556A-10

Query Match 66.4%; Score 14.6; DB 3; Length 30;
Best Local Similarity 81.0%; Pred. No. 3.9e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACCCTCAGCCACAT 21
| | | | | | | | | | | | | | | | | | | | | |
DB 5 CACCTTCACTCCACCCCAT 25

RESULT 5
US-08-897-556A-38
; Sequence 38, Application US/08897556A
; Patent No. 6570062
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA J.
; TITLE OF INVENTION: SYNTHETIC GENES FOR PLANT GUMS AND OTHER
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,556A
; FILING DATE: 21-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837

```

? REFERENCE/DOCKET NUMBER: CHU-029080
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 705-8410
? TELEFAX: (415) 397-8338
? INFORMATION FOR SEQ ID NO: 38:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 30 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: linear
? MOLECULE TYPE: other nucleic acid
? DESCRIPTION: /desc = "DNA"
US-08-897-556A-38

```

Query Match	66.4%	Score 14.6	DB 3	Length 30
Best Local Similarity	81.0%	Pred. No. 3.9e+03		
Matches 17, Conservative	0	Mismatches 4	Indels 0	Gaps 0

```

Qy      1 CTCCTCACCTCCAGCCACAT 21
         | ||| ||||| || |||
Db      5 CACCTTCACTCCACCCCAT 25

```

```

RESULT 6
US-09-547-693-10
: Sequence 10, Application US/09547693
: Patent No. 6639050
: GENERAL INFORMATION:
: APPLICANT: Kieliszewski, Marcia
: TITLE OF INVENTION: Synthetic Genes for Plant Gums and Other Hydroxyproline-Rich
: TITLE OF INVENTION: Glycoproteins
: FILE REFERENCE: OHU-04089
: CURRENT APPLICATION NUMBER: US/09/547,693
: CURRENT FILING DATE: 2000-04-12
: NUMBER OF SEQ ID NOS: 236
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 10
: LENGTH: 30
: TYPE: DNA
: ORGANISM: Artificial/Unknown
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Synthetic
: US-09-547-693-10

```

Query Match	66.4%	Score 14.6	DB 3	Length 30
Best Local Similarity	81.0%	Pred. No. 3.9e+03		
Matches 17; Conservative	0	Mismatches 4	Indels 0	Gaps 0

```

qy      1 CTCCTCACCTCGAGCCACAT 21
          | ||| ||||| |||
db      5 CACCTTCACCTCGACCCCGAT 25

```

```

? RESULT 7
? US-09-547-693--38
? Sequence 38 Application US/09547693
? Patent No. 6639050
? GENERAL INFORMATION:
? APPLICANT: Kiciłiszewski, Marcia
? TITLE OF INVENTION: Synthetic Genes for Plant Gums and Other Hydroxyproline-Rich
? FILE REFERENCE: OHU-04089
? CURRENT APPLICATION NUMBER: US/09/547,693
? CURRENT FILING DATE: 2000-04-12
? NUMBER OF SEQ ID NOS: 236
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 38
? LENGTH: 30
? TYPE: DNA
? ORGANISM: Artificial/Unknown
? FEATURE:
? NAME/KEY: misc_feature

```

OTHER INFORMATION: Synthetic
US-09-547-693-38

Query Match	66.4%	Score 14.6;	DB 3;	Length 30;
Best Local Similarity	81.0%	Pred. No. 3.9e+03;		
Matches 17;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0

```

QY      1 CTCCCTCACCCTCCAGCCACAT 21
          |||||
Db      5 CACCTTCACTCCACCCCAT 25

```

```

RESULT 8
US-08-529-1908-14/C
; Sequence 14, Application US/08529190B
; Patent No. 5833991
; GENERAL INFORMATION:
; APPLICANT: Masucci, Maria G.
; TITLE OF INVENTION: GLYCINE-CONTAINING SEQUENCES
; TITLE OF INVENTION: CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,190B
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE9501324-9
; FILING DATE: 10-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US08/522,595
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen A
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 3255/53015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; US-08-529-1908-14

Query Match 64.5%; Score 14.2; DB 2; Length 24;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

0Y 3 CCCTCACTTCAGGCACAT 21
   ||| ||||| |||
Db 23 CCGCACCTTCAGCTCAT 5

RESULT 9
US-07-792-600-8
; Sequence 8, Application US/07792600
; Patent No. 6008045
; GENERAL INFORMATION:

```

Query Match	64.5%	Score 14.2;	DB 2;	Length 24;
Best Local Similarity	84.2%	Pred. NO. 5.6e+03;		
Matches 16; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy	3	CCCTCACCCTCAGCCACAT	21
Db	23	CCCGCACCTCAGCTTCAT	5

RESULT 9
US-07-792-600-8
; Sequence 8, Application US/07792600
; Patent No. 6008045
; GENERAL INFORMATION:

APPLICANT: COPELAND, WILLIAM C.
APPLICANT: WANG, TERESA S.-P.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: TEMPLATE-DEPENDENT ENZYMATIC
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter G. Carroll
STREET: 220 Montgomery Street, Suite 710
CITY: San Francisco
STATE: California
COUNTRY: U.S.A.
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/792,600
FILING DATE: 1991.11.15
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: STDU-00097
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-792-600-8

Query Match	64.5%;	Score 14.2;	DB 3;	Length 24;
Best Local Similarity	84.2%;	Pred. No. 5.6e+03;		
Matches 16;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps				

OY	4	CCTCACCCTCCAGCCACATG	22
Db	1	CTTCACCTCCAGCCAGTG	19

RESULT 10
US-09-157-021-8
; Sequence 8, Application US/09157021A
; Patent No. 6100023
; GENERAL INFORMATION:
; APPLICANT: Copeland, William C.
; APPLICANT: Wang, Teresa S. F.
; TITLE OF INVENTION: Drug Design Assay
; FILE REFERENCE: STDU-03484
; CURRENT APPLICATION NUMBER: US/09/157,021A
; CURRENT FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: 07/792,600
; EARLIER FILING DATE: 1991-11-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-157-021-8

Query Match	64.5%;	Score 14.2;	DB 3;	Length 24;
Best Local Similarity	84.2%;	Pred. No. 5.6e+03;		
Matches 16;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps				

OY	4	CCTCACCCTCCAGCCACATG	22
----	---	----------------------	----

```

Db      1 CTTACCTCCAGCGCAGGTG   19          |||||
RESULT 11 US-09-156-842-8
; Sequence 8 , Application US/09156842A
Patent No. 6103473
GENERAL INFORMATION:
APPLICANT: Copeland, William C.
APPLICANT: Wang, Teresa S. F.
TITLE OF INVENTION: Drug Screening
FILE REFERENCE: STDU-03485
CURRENT APPLICATION NUMBER: US/09/156,842A
CURRENT FILING DATE: 1998-09-18
EARLIER APPLICATION NUMBER: 07/792,600
EARLIER FILING DATE: 1991-11-15
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-156-842-8
Query Match           64.5%; Score 14.2; DB 3; Length 24;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches    16; Conservative     0; Mismatches       3; Indels         0; Gaps        0;
```

```

RESULT 12
US-09-591-514-8
; Sequence 8, Application US/09591514
; Patent No. 6670161
; GENERAL INFORMATION:
; APPLICANT: Copeland, William C.
; APPLICANT: Wang, Teresa S. F.
; TITLE OF INVENTION: Drug Design Assay
; FILE REFERENCE: STDU-03464
; CURRENT APPLICATION NUMBER: US/09/591,514
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: US/09/157,021
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 07/792,600
; PRIOR FILING DATE: 1991-11-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-591-514-8

Query Match          64.5%; Score 14.2; DB 3; Length 24;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy          4 CCTCACTCCAGCCACATG 22
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
DB          1 CTTACCTCCAGCCAGGTG 19

RESULT 13
US-09-887-145-6/C
; Sequence 6, Application US/09887145

```

```
Patent No. 6780641
GENERAL INFORMATION:
APPLICANT: Kim, Seung U
TITLE OF INVENTION: Immobilized human microglia
cell and continuous cell line
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: David Praehker, Esq.
STREET: P.O. Box 5387
CITY: Magnolia
STATE: Massachusetts
COUNTRY: USA
ZIP: 01930
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage
COMPUTER: Dell PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Microsoft Word version 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/887,145
FILING DATE: 22-Jun-2001
CLASSIFICATION: Unknown
ATTORNEY/AGENT INFORMATION:
NAME: David Praehker, Esq.
REGISTRATION NUMBER: 29,693
REFERENCE/DOCKET NUMBER: UBC-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (978) 525-3794
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-887-145-6

Query Match          63.6%; Score 14; DB 3; Length 24;
Best Local Similarity 77.3%; Pred. No. 6.8e+03;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCAGCCATG 22
Db 22 CACCTCAGCGCATCAATG 1

RESULT 14
US-09-674-824-5/c
Sequence 5, Application US/09674824
Patent No. 6890732
GENERAL INFORMATION:
APPLICANT: Lotz, et al
TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH CODE FOR ENZYMES DERIVED FROM WHEAT
FILE REFERENCE: AGR 1998/M 205; FLH514413-3848
CURRENT APPLICATION NUMBER: US/09/674,824
CURRENT FILING DATE: 2000-02-16
PRIOR APPLICATION NUMBER: PCT/EP99/03156
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: DE 198 20607.0
PRIOR FILING DATE: 1998-05-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence, primer
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)..(20)
OTHER INFORMATION: primer
US-09-674-824-5
```

```
Query Match          62.7%; Score 13.8; DB 3; Length 20;
Best Local Similarity 88.2%; Pred. No. 8.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CTCACCTCCAGCCACA 20
Db 17 CTCACCTCCAGCCACA 1

RESULT 15
US-09-487-368A-216
Sequence 216, Application US/09487368A
Patent No. 6261840
GENERAL INFORMATION:
APPLICANT: Lex M. Cowseart
APPLICANT: Jacqueline Wyatt
TITLE OF INVENTION: ANTISENSE MODULATION OF PTP1B EXPRESSION
FILE REFERENCE: RTS-0093
CURRENT APPLICATION NUMBER: US/09/487,368A
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 240
SEQ ID NO 216
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Antisense Oligonucleotide
US-09-487-368A-216

Query Match          61.8%; Score 13.6; DB 3; Length 20;
Best Local Similarity 80.0%; Pred. No. 9.8e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCAGCCACA 20
Db 1 CTGCCCCAGCAGCCAGCCACA 20
```

Search completed: March 25, 2006, 13:03:00
Job time : 130.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using BW model

Run on: March 25, 2006, 13:59:41 ; Search time 660.5 Seconds

(without alignments)
275.437 Million cell updates/sec

Title: US-10-798-652-6

Perfect score: 22

Sequence: 1 CTCCTCACCCTCCAGCCACATG 22

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 10535742

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	US-10-798-652-6	Sequence 6, Appli
2	20.4	92.7	22	US-10-798-652-7	Sequence 7, Appli
3	15.8	71.8	25	US-10-719-900-593282	Sequence 693282,
4	15.8	71.8	25	US-10-956-157-252303	Sequence 252303,
5	15.6	70.9	25	US-10-485-048-86	Sequence 86, Appli
6	15.6	70.9	25	US-11-036-317-179644	Sequence 179644,
7	15.6	70.9	25	US-11-036-317-250940	Sequence 250940,
8	15.6	70.9	25	US-11-036-317-253919	Sequence 253919,
9	15.6	70.9	25	US-11-036-317-268615	Sequence 268615,
10	15.6	70.9	25	US-11-036-317-362146	Sequence 362146,
11	15.6	70.9	25	US-11-036-317-391184	Sequence 391184,
12	15.4	70.0	25	US-11-036-317-579816	Sequence 579816,
13	15.2	69.1	25	US-10-215-112-12019	Sequence 12019, A
14	15.2	69.1	25	US-10-215-112-12145	Sequence 12145, A
15	15.2	69.1	25	US-11-036-317-278200	Sequence 278200,
16	15.2	69.1	25	US-11-036-317-386177	Sequence 386177,
17	14.8	67.3	25	US-10-719-900-217185	Sequence 217185,
18	14.8	67.3	25	US-10-719-900-217186	Sequence 217186,
19	14.8	67.3	25	US-10-956-157-132785	Sequence 132785,
20	14.6	66.4	25	US-10-719-956-459914	Sequence 459914,
21	14.6	66.4	25	US-11-036-317-280075	Sequence 280075,
22	14.6	66.4	25	US-11-036-317-360952	Sequence 360952,
23	14.6	66.4	25	US-11-036-317-469308	Sequence 469308,

24	14.6	66.4	25	10	US-11-036-317-516411	Sequence 516411,
C 25	14.6	66.4	25	10	US-11-036-317-897472	Sequence 897472,
C 26	14.6	66.4	25	10	US-11-036-317-927798	Sequence 927798,
C 27	14.6	66.4	25	10	US-11-036-317-980808	Sequence 980808,
C 28	14.6	66.4	25	10	US-11-060-756-206500	Sequence 206500, A
29	14.6	66.4	25	10	US-11-060-756-206518	Sequence 206518, A
30	14.6	66.4	30	6	US-10-437-708-10	Sequence 10, Appli
31	14.6	66.4	30	6	US-10-437-708-38	Sequence 38, Appli
32	14.6	66.4	30	6	US-10-395-402-10	Sequence 10, Appli
33	14.6	66.4	30	6	US-10-395-402-38	Sequence 38, Appli
34	14.6	66.4	30	8	US-10-257-199-10	Sequence 10, Appli
35	14.6	66.4	30	8	US-10-257-199-38	Sequence 38, Appli
36	14.6	66.4	30	9	US-10-418-032-10	Sequence 10, Appli
37	14.6	66.4	30	9	US-10-418-032-38	Sequence 38, Appli
38	14.2	64.5	25	7	US-10-681-773-98599	Sequence 98599, A
39	14.2	64.5	25	7	US-10-681-773-98599	Sequence 98599, A
C 40	14.2	64.5	25	8	US-10-719-900-221421	Sequence 221421,
C 41	14.2	64.5	25	8	US-10-719-900-693283	Sequence 693283,
42	14.2	64.5	25	9	US-10-956-157-70581	Sequence 70581, A
43	14.2	64.5	25	9	US-10-956-157-70582	Sequence 70582, A
44	14.2	64.5	25	9	US-10-956-157-70583	Sequence 70583, A
45	14.2	64.5	25	9	US-10-956-157-70584	Sequence 70584, A

ALIGNMENTS

```
RESULT 1
US-10-798-652-6
; Sequence 6, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICANT: Yonjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; FILE REFERENCE: 3382-POJ136US01
; CURRENT APPLICATION NUMBER: US/10/798,652
; PRIOR FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/455,698
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: probe
US-10-798-652-6

Query Match      100.0%; Score 22; DB 8; Length 22;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTCCTCACCCTCCAGCCACATG 22
Db       1 CTCCTCACCCTCCAGCCACATG 22

RESULT 2
US-10-798-652-7
; Sequence 7, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICANT: Yonjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; FILE REFERENCE: 3382-POJ136US01
; CURRENT APPLICATION NUMBER: US/10/798,652
; PRIOR FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/455,698
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 7
```

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 7
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: probe
US-10-798-652-7

Query Match 92.7% Score 20.4; DB 8; Length 22;
Best Local Similarity 95.5%; Pred. No. 53;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCACTCCAGCCACATG 22
DB 1 CTCCTCGCCTCCAGCCACATG 22

RESULT 3
US-10-719-900-693282/c
Sequence 693282, Application US/10719900
Publication No. US20050026164A1

GENERAL INFORMATION:
APPLICANT: Xue Mei Zhou
TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
FILE REFERENCE: 3528.1
CURRENT APPLICATION NUMBER: US/10/719,900
CURRENT FILING DATE: 2003-11-20
PRIOR APPLICATION NUMBER: 60/427,808
PRIOR FILING DATE: 2002-11-20
NUMBER OF SEQ ID NOS: 982914
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 693282
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-10-719-900-693282

Query Match 71.8% Score 15.8; DB 8; Length 25;
Best Local Similarity 89.5%; Pred. No. 4.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCCTCACTCCAGCCACA 20
DB 22 TCCCTCACTCCAGCCACA 4

RESULT 4
US-10-956-157-252303/c
Sequence 252303, Application US/10956157
Publication No. US20050118625A1

GENERAL INFORMATION:
APPLICANT: Myeth
APPLICANT: Mounts, William
TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
FILE REFERENCE: 031896-043000 (AM 101081)
CURRENT APPLICATION NUMBER: US/10/956,157
CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 319805
SOFTWARE: Patent in version 3.2
SEQ ID NO 252303
LENGTH: 25
TYPE: DNA
ORGANISM: Probe Sequence
US-10-956-157-252303

Query Match 71.8% Score 15.8; DB 9; Length 25;
Best Local Similarity 89.5%; Pred. No. 4.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCACTCCAGCCAC 19
DB 1 CTCCTCACTCCAGCCAC 19

DB 24 CTCCTCACTCCAGCCAC 6

RESULT 5
US-10-485-048-86/c
Sequence 86, Application US/10485048
Publication No. US20050058657A1
GENERAL INFORMATION:
APPLICANT: ERTL, Peter Franz
APPLICANT: TITE, John Philip
APPLICANT: VAN WELY, Catherine Anne

APPLICANT: VOSS, Gerald
TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
FILE REFERENCE: B45279
CURRENT APPLICATION NUMBER: US/10/485,048
CURRENT FILING DATE: 2004-01-26
PRIOR APPLICATION NUMBER: PCT/EP02/08343
PRIOR FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: GB 0118367.2
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 98
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 86
LENGTH: 25
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: primer (Figure 21)
US-10-485-048-86

Query Match 70.9% Score 15.6; DB 9; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACTCCAGCCACATG 22
DB 22 CTCCTCACTCCAGCCACATG 1

RESULT 6
US-11-036-317-179644/c
Sequence 179644, Application US/11036317
Publication No. US20050214823A1

GENERAL INFORMATION:
APPLICANT: Blume, John
TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
FILE REFERENCE: 3654.1
CURRENT APPLICATION NUMBER: US/11/036,317
CURRENT FILING DATE: 2005-01-13
PRIOR APPLICATION NUMBER: US 60/536,639
PRIOR FILING DATE: 2004-01-13
NUMBER OF SEQ ID NOS: 991174
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 179644
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-11-036-317-179644

Query Match 70.9% Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCACTCCAGCCACATG 22
DB 23 CCCACTCCCTCCAGCCACAG 2

RESULT 7
US-11-036-317-250940
Sequence 250940, Application US/11036317


```
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 250940
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-250940

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 2 CTCCTCCCATGAGCCACCTG 23

RESULT 8
US-11-036-317-253919/c
/ Sequence 253919, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 253919
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-253919

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 22 CCCACTCCCTCCAGCCACAG 1

RESULT 9
US-11-036-317-268615
/ Sequence 268615, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 268615
```

```
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-268615

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 1 CTCCTCCCATGAGCCACCTG 22

RESULT 10
US-11-036-317-362146
/ Sequence 362146, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 362146
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-362146

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 3 CTCCTCCCATGAGCCACCTG 24

RESULT 11
US-11-036-317-391184
/ Sequence 391184, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 391184
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-391184

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.6e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCAGCTCCAGCCACATG 22
    |||||
Db 4 CTCCTCCCATGAGCCACCTG 25
```

```
RESULT 12
US-11-036-317-579816/c
; Sequence 579816, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; PRIOR FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO: 579816
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-579816

Query Match          70.0%; Score 15.4; DB 10; Length 25;
Best Local Similarity 94.1%; Pred. No. 6.8e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2  TCCTCAGCTCCAGCCCA 18
Db      19 TCCTCAGCTCCAGCCCA 3

RESULT 13
US-10-215-112-12019
; Sequence 12019, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; PRIOR FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12019
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-12019

Query Match          69.1%; Score 15.2; DB 5; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3  CCCTCAGCTCCAGCCCATG 22
Db      1  CCTTCAGTCCACACCATG 20

RESULT 14
US-10-215-112-12145
; Sequence 12145, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; PRIOR FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12145
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-12145

Query Match          69.1%; Score 15.2; DB 5; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3  CCCTCAGCTCCAGCCCATG 22
Db      1  CCTTCAGTCCATCCACCATG 20

RESULT 15
US-11-036-317-278200/c
; Sequence 278200, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; PRIOR FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO: 278200
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-278200

Query Match          69.1%; Score 15.2; DB 10; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  CTCCTCAGCTCCAGCCACA 20
Db      21 CCAGTCCCTCCAGCCACA 2

Search completed: March 25, 2006, 15:50:31
Job time : 661.5 secs
```



```

; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 233742
; LENGTH: 24
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-233742
```

```
Query Match
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 CCCTCAGCTCCAGCCACATG 22
    |||:|||||:|||||:|
Db 1 CCCCAACCCGCGCCACAG 20
```

```
RESULT 3
US-10-310-914A-668834/c
; Sequence 668834, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shlier, Kyuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 668834
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-668834
```

```
Query Match
Best Local Similarity 74.5%; Score 16.4; DB 8; Length 19;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 CCCTCAGCTCCAGCCACA 20
    |||:|||||:|||||:|
Db 18 CCCCACCTCCAGCCACA 1
```

```
RESULT 4
US-10-310-914A-317759
; Sequence 317759, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shlier, Kyuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 317759
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-317759
```

```
Query Match
Best Local Similarity 74.5%; Score 16.4; DB 8; Length 20;
Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 4 CCTCAGCTCCAGCCACAT 21
    ||:||||:|||||:|
Db 3 CCUACCCUCCGCGCACAU 20
```

```
RESULT 5
US-10-310-914A-377554/c
; Sequence 377554, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shlier, Kyuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 377554
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-377554
```

```
Query Match
Best Local Similarity 74.5%; Score 16.4; DB 8; Length 20;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCAGCTCCAGCCA 18
    |||:|||||:|||||:|
Db 18 CCCCCTCAGCTCCAGCCA 1
```

```
RESULT 6
US-10-310-914A-91862/c
; Sequence 91862, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shlier, Kyuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 91862
; LENGTH: 25
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-91862
```

```
Query Match
Best Local Similarity 73.6%; Score 16.2; DB 8; Length 25;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 TCCTCAGCTCCAGCCATG 22
    |||:|||||:|||||:|
Db 23 TCCTTAAGCTCCAGCTCTG 3
```

```
RESULT 7
US-10-310-914A-91863/c
; Sequence 91863, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shlier, Kyuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
```

SOFTWARE: PatentIn version 3.3
SEQ ID NO 91863
LENGTH: 25
TYPE: RNA
ORGANISM: Human
US-10-310-914A-91863

Query Match 73.6%; Score 16.2; DB 8; Length 25;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TCCTCAGCTCCAGCCACATG 22
DB 23 TCCTTAAGCTCCAGCTCCTG 3

RESULT 8
US-10-310-914A-448406
Sequence 448406, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kvyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 448406
LENGTH: 20
TYPE: RNA
ORGANISM: Human
US-10-310-914A-448406

Query Match 71.8%; Score 15.8; DB 8; Length 20;
Best Local Similarity 78.9%; Pred. No. 2.3e+03;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCAC 19
DB 1 CUCCGACACUCCAGGCAC 19

RESULT 9
US-10-310-914A-724597/c
Sequence 724597, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kvyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 724597
LENGTH: 21
TYPE: RNA
ORGANISM: Human
US-10-310-914A-724597

Query Match 71.8%; Score 15.8; DB 8; Length 21;
Best Local Similarity 89.5%; Pred. No. 2.3e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCTCAGCTCCAGCCACA 20
DB 19 TCCTCATCTCTGCGACA 1

RESULT 10
US-10-310-914A-448407
Sequence 448407, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kvyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 448407
LENGTH: 23
TYPE: RNA
ORGANISM: Human
US-10-310-914A-448407

Query Match 71.8%; Score 15.8; DB 8; Length 23;
Best Local Similarity 78.9%; Pred. No. 2.4e+03;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCAC 19
DB 1 CUCCGACACUCCAGGCAC 19

RESULT 11
US-10-310-914A-1002544
Sequence 1002544, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kvyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 1002544
LENGTH: 23
TYPE: RNA
ORGANISM: Human
US-10-310-914A-1002544

Query Match 71.8%; Score 15.8; DB 8; Length 23;
Best Local Similarity 78.9%; Pred. No. 2.4e+03;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCAC 19
DB 4 CUCCGACACUCCAGCCCC 22

RESULT 12
US-10-310-914A-180307/c
Sequence 180307, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kvyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3

SEQ ID NO 180307
LENGTH: 24
TYPE: RNA
ORGANISM: Human
US-10-310-914A-180307

Query Match 71.8%; Score 15.8; DB 8; Length 24;
Best Local Similarity 89.5%; Pred. No. 2.4e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTCAGCTCCAGCCACATG 22
DB 24 CCTCACTCCAGCCACCTG 6

RESULT 13
US-10-310-914A-66290
Sequence 66290, Application US/10310914A
Publication No. US2006000332A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shlier, Kuzat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 66290
LENGTH: 24
TYPE: RNA
ORGANISM: Human
US-10-310-914A-66290

Query Match 70.9%; Score 15.6; DB 8; Length 24;
Best Local Similarity 72.7%; Pred. No. 2.9e+03;
Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCCACATG 22
DB 2 CACCCCTCCAGCCACATG 23

RESULT 14
US-11-101-244-196424/C
Sequence 196424, Application US/11101244
Publication No. US20050246794A1
GENERAL INFORMATION:
APPLICANT: Dharmoon, Inc.
APPLICANT: Khvorova, Anastasia
APPLICANT: Reynolds, Angela
APPLICANT: Leake, Devin
APPLICANT: Marshall, William
APPLICANT: Scaringe, Stephen
TITLE OF INVENTION: Functional and Hyperfunctional siRNA
FILE REFERENCE: 13499US
CURRENT APPLICATION NUMBER: US/11/101,244
CURRENT FILING DATE: 2005-04-07
PRIOR APPLICATION NUMBER: 60/502,050
PRIOR FILING DATE: 2003-09-10
PRIOR APPLICATION NUMBER: 60/426,137
PRIOR FILING DATE: 2002-11-14
NUMBER OF SEQ ID NOS: 1591911
SOFTWARE: Proprietary
SEQ ID NO 196424
LENGTH: 19
TYPE: RNA
ORGANISM: Homo sapiens
US-11-101-244-196424

Query Match 70.0%; Score 15.4; DB 12; Length 19;
Best Local Similarity 94.1%; Pred. No. 3.4e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCCTCAGCTCCAGCC 17
DB 17 CTCCTCAGCTCCAGCC 1

RESULT 15
US-11-083-784-196424/C
Sequence 196424, Application US/11083784
Publication No. US20050245475A1
GENERAL INFORMATION:
APPLICANT: Dharmoon, Inc.
APPLICANT: Khvorova, Anastasia
APPLICANT: Reynolds, Angela
APPLICANT: Leake, Devin
APPLICANT: Marshall, William
APPLICANT: Scaringe, Stephen
TITLE OF INVENTION: Functional and Hyperfunctional siRNA
FILE REFERENCE: 13499US
CURRENT APPLICATION NUMBER: US/11/083,784
CURRENT FILING DATE: 2005-03-18
PRIOR APPLICATION NUMBER: US/10/714,333
PRIOR FILING DATE: 2003-11-14
PRIOR APPLICATION NUMBER: 60/502,050
PRIOR FILING DATE: 2003-09-10
PRIOR APPLICATION NUMBER: 60/426,137
PRIOR FILING DATE: 2002-11-14
NUMBER OF SEQ ID NOS: 1591911
SOFTWARE: Proprietary
SEQ ID NO 196424
LENGTH: 19
TYPE: RNA
ORGANISM: Homo sapiens
US-11-083-784-196424

Query Match 70.0%; Score 15.4; DB 13; Length 19;
Best Local Similarity 94.1%; Pred. No. 3.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTCCAGCC 17
DB 17 CTCCTCAGCTCCAGCC 1

Search completed: March 25, 2006, 14:54:04
Job time : 315.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 13:42:31 ; Search time 959 Seconds
(without alignments)
1304.021 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22
Sequence: 1 ctcctcgctccagccacatg 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 1641224

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GenEmbl:

1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.2	64.5	30	6	A79795 Sequence 3
2	14	63.6	25	6	AX701087 Sequence
3	13.8	62.7	20	6	CS109308 Sequence
4	13.8	62.7	24	6	AX444501 Sequence
5	13.6	61.8	22	6	BD086268 Vector. 8
6	13.6	61.8	22	6	AX003404 Sequence
7	13.6	61.8	27	6	AR091427 Sequence
8	13.6	61.8	27	6	AR091430 Sequence
9	13.6	61.8	27	6	AR125632 Sequence
10	13.6	61.8	27	6	AR125635 Sequence
11	13.6	61.8	27	6	AR125635 Sequence
12	13	59.1	23	6	AX732524 Sequence
13	13	59.1	24	6	AR073296 Sequence
14	13	59.1	24	6	AR052984 Sequence
15	13	59.1	25	6	AR076350 Sequence
16	13	59.1	30	6	AR306619 Sequence
17	13	59.1	30	6	AR306628 Sequence
18	13	59.1	30	6	AR340054 Sequence

19	13	59.1	30	6	AR340063 Sequence
20	13	59.1	30	6	AR412123 Sequence
21	13	59.1	30	6	AR412132 Sequence
22	12.8	58.2	17	6	AR379426 Sequence
23	12.8	58.2	20	6	AX598365 Sequence
24	12.8	58.2	21	6	BD174200 Sequence
25	12.8	58.2	21	6	BD185155 Sequence
26	12.8	58.2	21	6	CS012518 Sequence
27	12.8	58.2	22	6	CS012518 Sequence
28	12.6	57.3	19	6	CS095314 Sequence
29	12.6	57.3	19	6	CS095413 Sequence
30	12.6	57.3	22	6	BD249866 Sequence
31	12.6	57.3	22	6	AX033501 Sequence
32	12.6	57.3	24	6	A57517 Sequence 9
33	12.6	57.3	24	6	A57536 Sequence 28
34	12.6	57.3	24	6	AR052983 Sequence
35	12.6	57.3	24	6	AR096578 Sequence
36	12.6	57.3	24	6	AR105954 Sequence
37	12.6	57.3	24	6	AR442798 Sequence
38	12.6	57.3	24	6	AX290630 Sequence
39	12.6	57.3	25	6	BD174183 Sequence
40	12.6	57.3	25	6	BD185138 Cell diff
41	12.6	57.3	25	6	AX501644 Sequence
42	12.6	57.3	25	6	AX501645 Sequence
43	12.6	57.3	25	6	AX501646 Sequence
44	12.6	57.3	25	6	AX501647 Sequence
45	12.6	57.3	25	6	AX501648 Sequence

ALIGNMENTS

RESULT 1	A79795	Sequence 3 from Patent WO9714795.	30 bp	DNA	linear	PAT 20-OCT-1999
LOCUS	A79795	Sequence 3 from Patent WO9714795.	30 bp	DNA	linear	PAT 20-OCT-1999
DEFINITION	A79795	Sequence 3 from Patent WO9714795.	30 bp	DNA	linear	PAT 20-OCT-1999
ACCESSION	A79795	Sequence 3 from Patent WO9714795.	30 bp	DNA	linear	PAT 20-OCT-1999
VERSION	A79795.1	GI:6092713	30 bp	DNA	linear	PAT 20-OCT-1999
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
ORIGIN						
Query Match	64.5%	Score 14.2;	DB 6;	Length 30;		
Best local Similarity	84.2%	Pred. No. 1.4e+05;				
Matches	16;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
DB	11	CCGCGCTCCAGCTCATG 29				
RESULT 2	AX701087/c	Sequence 86 from Patent WO03011334.	25 bp	DNA	linear	PAT 03-APR-2003
LOCUS	AX701087	Sequence 86 from Patent WO03011334.	25 bp	DNA	linear	PAT 03-APR-2003
DEFINITION	AX701087	Sequence 86 from Patent WO03011334.	25 bp	DNA	linear	PAT 03-APR-2003
ACCESSION	AX701087	Sequence 86 from Patent WO03011334.	25 bp	DNA	linear	PAT 03-APR-2003
VERSION	AX701087.1	GI:29536869	25 bp	DNA	linear	PAT 03-APR-2003
KEYWORDS						
SOURCE						
synthetic construct						

ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE
1 Ertl, P.F., Tite, J.P., van Wely, C.A. and Voss, G.
TITLE Vaccine comprising gp120 and nef and/or tat for the immunisation
JOURNAL against hiv
Patent: WO 03011334-A 86 13-FEB-2003;
GlaXoSmithKline Biologicals S.A. (BE) ; GLAXO GROUP LIMITED (GB)
LOCATION/Qualifiers

FEATURES
source
1..25
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="primer (figure 21)"

ORIGIN

Query Match 63.6%; Score 14; DB 6; Length 25;
Best Local Similarity 77.3%; Pred. No. 1.7e+05;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy
1 CTCGCTCGCCTCAGCCACATG 22
22 CTTTCCACCTCTCTGACATG 1

Db
1 CTCGCTCGCCTCAGCCACATG 22
22 CTTTCCACCTCTCTGACATG 1

RESULT 3
CS109308 20 bp DNA linear PAT 22-JUN-2005
LOCUS CS109308
DEFINITION Sequence 7 from Patent WO2005052140.
ACCESSION CS109308
VERSION CS109308.1 GI:68148122
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Carcharhini;
Hominidae; Homo.
REFERENCE
1 Shi, S., Seo, B.M. and Miura, M.
AUTHORS Multipotent postnatal stem cells from human periodontal ligament
TITLE and uses thereof
JOURNAL Patent: WO 2005052140-A 7 09-JUN-2005;
Department of Health and Human Services (US)
LOCATION/Qualifiers
FEATURES
source
1..20
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 62.7%; Score 13.8; DB 6; Length 20;
Best Local Similarity 88.2%; Pred. No. 2.1e+05;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy
5 CTCGCTCGCCTCAGCCACAT 21
1 CTCGCTCGCCTCAGCCACAT 17

Db
1 CTCGCTCGCCTCAGCCACAT 17

RESULT 4
AX444501 24 bp DNA linear PAT 03-JUL-2002
LOCUS AX444501
DEFINITION Sequence 956 from Patent WO0216649.
ACCESSION AX444501
VERSION AX444501.1 GI:21691779
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Gunderson, K.
AUTHORS Probes and decoder oligonucleotides
TITLE Patent: WO 0216649-A 956 28-FEB-2002;
JOURNAL

ILLUMINA, Inc. (US)
LOCATION/Qualifiers

FEATURES
source
1..24
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Computer Generated Probe Sequence."

ORIGIN

Query Match 62.7%; Score 13.8; DB 6; Length 24;
Best Local Similarity 88.2%; Pred. No. 2.1e+05;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy
5 CTCGCTCGCCTCAGCCACAT 21
21 CTCGCTCGCCTCAGCCACAT 5

Db
1 CTCGCTCGCCTCAGCCACAT 5

RESULT 5
BD086268 22 bp DNA linear PAT 27-AUG-2002
LOCUS BD086268
DEFINITION Vector.
ACCESSION BD086268
VERSION BD086268.1 GI:22631878
KEYWORDS
SOURCE JP 2001525168-A/21.
ORGANISM JP 2001525168-A/21.
synthetic construct
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 22)
AUTHORS Shen, S., Schedl, A. and Harnmar, A.J.
TITLE Vector
JOURNAL Patent: JP 2001525168-A 21 11-DEC-2001;
MEDICAL RESEARCH COUNCIL
COMMENT OS Artificial Sequence
PN JP 2001525168-A/21
PD 11-DEC-2001 JP 2000523326
PF 27-NOV-1998 JP 2000523326
PR 28-NOV-1997 GB 9725311.6, 28-NOV-1997 GB 9725313.2 PR
20-MAR-1998 GB 9806072.6, 05-NOV-1998 GB 9824275.3 PI
SANBING SHEN, ANDREAS SCHEDL, ANTHONY JOHN HARNMAR PC
C12N15/09, C12N15/00
CC Description of Artificial Sequence: Primer
FH Key
FT source
1..22
LOCATION/Qualifiers
FEATURES
source
1..22
/organism="Artificial Sequence".
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 61.8%; Score 13.6; DB 6; Length 22;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy
2 TCCCTCGCCTCAGCCACAT 21
1 TTCATCACCCTCATCCACAT 20

Db
1 TTCATCACCCTCATCCACAT 20

RESULT 6
AX003404 22 bp DNA linear PAT 24-AUG-2000
LOCUS AX003404
DEFINITION Sequence 21 from Patent WO9928449.
ACCESSION AX003404
VERSION AX003404.1 GI:9927208
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Schedl, A. and Harnmar, A.J.

TITLE Vectors
JOURNAL Patent: WO 9928449-A 21 10-JUN-1999;
SCHEDL ANDREAS (DE); HARMAR ANTHONY JOHN (GB)
FEATURES
SOURCE Location/Qualifiers
1..22
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Primer"

ORIGIN

Query Match 61.8%; Score 13.6; DB 6; Length 22;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TCCCTCGCTCCAGCCACAT 21
Db 1 TTCAATCACCCTCCATCCACAT 20

RESULT 7

AR091427/c
LOCUS AR091427 27 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 19 from patent US 5994109.
ACCESSION AR091427
VERSION AR091427.1 GI:10018182
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 27)
AUTHORS Woo,S.L.C., Smith,L.C., Cristiano,R.J., Gottchalk,S. and Sparrow,J.
TITLE Nucleic acid transporter system and methods of use
JOURNAL Patent: US 5994109-A 19 30-NOV-1999;
FEATURES Location/Qualifiers
1..27
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 61.8%; Score 13.6; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

RESULT 8
AR091430/c
LOCUS AR091430 27 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 22 from patent US 5994109.
ACCESSION AR091430
VERSION AR091430.1 GI:10018185
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 27)
AUTHORS Woo,S.L.C., Smith,L.C., Cristiano,R.J., Gottchalk,S. and Sparrow,J.
TITLE Nucleic acid transporter system and methods of use
JOURNAL Patent: US 5994109-A 22 30-NOV-1999;
FEATURES Location/Qualifiers
1..27
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 61.8%; Score 13.6; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

RESULT 9
AR125632/c
LOCUS AR125632 27 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 19 from patent US 6177554.
ACCESSION AR125632
VERSION AR125632.1 GI:14111694
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 27)
AUTHORS Woo,S.L.C., Smith,L.C., Cristiano,R.J., Gottchalk,S. and Sparrow,J.
TITLE Nucleic acid transporter systems
JOURNAL Patent: US 6177554-A 19 23-JAN-2001;
FEATURES Location/Qualifiers
1..27
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 61.8%; Score 13.6; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

RESULT 10
AR125635/c
LOCUS AR125635 27 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 22 from patent US 6177554.
ACCESSION AR125635
VERSION AR125635.1 GI:14111697
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 27)
AUTHORS Woo,S.L.C., Smith,L.C., Cristiano,R.J., Gottchalk,S. and Sparrow,J.
TITLE Nucleic acid transporter systems
JOURNAL Patent: US 6177554-A 22 23-JAN-2001;
FEATURES Location/Qualifiers
1..27
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 61.8%; Score 13.6; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 2.6e+05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCTCCAGCCACA 20
Db 20 CTCCTCGCTCCAGCCACA 1

RESULT 11

AX732524
LOCUS AX732524 17 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 4158 from Patent WO03025175.
ACCESSION AX732524
VERSION AX732524.1 GI:30511867
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Search completed: March 25, 2006, 14:43:30
Job time : 960 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 13:22:55 ; Search time 372.5 Seconds
(without alignments)
393.619 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22

Sequence: 1 ctccctcgcctccagccacatg 22

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 4138570

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 0%
Listing first 45 summaries

Database :

N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005as:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	13	ADT97907 Human FGF
2	20.4	92.7	22	13	ADT97906 Human FGF
3	15.4	70.0	20	14	AA51146 Pyranair
4	14.8	67.3	20	14	ADW79585 Human ART
5	14.6	66.4	29	12	ADY00944 PCR prime
6	14.2	64.5	30	2	AAT90350 DNA encod
7	14	63.6	25	10	ADB67764 HIV-1 gp1
8	13.8	62.7	20	14	AA51123 Human c1a
9	13.8	62.7	24	6	ABQ00949 Oligonuc1
10	13.8	62.7	24	6	ABQ05868 Oligonuc1
11	13.8	62.7	24	6	ABQ05827 Oligonuc1
12	13.6	61.8	20	4	AA507411 PCR prime
13	13.6	61.8	22	2	AA57861 PCR prime
14	13.6	61.8	22	5	AAC85242 Reverse p
15	13.6	61.8	25	9	ACT188529 Human mic
16	13.6	61.8	25	9	ACH62883 DNA targ
17	13.6	61.8	25	9	ACH63009 DNA targ
18	13.6	61.8	27	3	AAA36522 Ligand se
19	13.6	61.8	27	3	AAA36549 DNA bindi

C	20	13.6	61.8	27	3	AAZ38194	AAZ38194 Sequence
C	21	13.6	61.8	27	3	AAZ39499	AAZ39499 C-myc pro
C	22	13.6	61.8	27	4	AAC82869	AAC82869 Nucleic a
C	23	13.6	61.8	27	4	AAC82872	AAC82872 Nucleic a
C	24	13.6	61.8	27	4	AA508479	AA508479 C-myc DNA
C	25	13.6	61.8	27	4	AA508482	AA508482 DNA 11gan
C	26	13.2	60.0	19	10	ADG35018	ADG35018 Human TNF
C	27	13.2	60.0	19	10	ADG34895	ADG34895 Human TNF
C	28	13.2	60.0	20	3	AAA53034	AAA53034 Human CDN
C	29	13.2	60.0	21	6	AB598432	AB598432 Human mul
C	30	13.2	60.0	22	10	ADD22516	ADD22516 Flathish
C	31	13.2	60.0	25	2	AAZ25239	AAZ25239 Maize cin
C	32	13.2	60.0	26	6	AD46882	AD46882 Probe H1P
C	33	13.2	60.0	26	14	ADY49468	ADY49468 Human neu
C	34	13	59.1	17	8	ABT38521	ABT38521 Tumour su
C	35	13	59.1	20	13	ADR87437	ADR87437 Epoxigena
C	36	13	59.1	20	14	ABE02351	ABE02351 Stokesia
C	37	13	59.1	22	10	ADC36189	ADC36189 Weed cont
C	38	13	59.1	22	13	ADR88302	ADR88302 Mus muscu
C	39	13	59.1	23	2	AAZ09311	AAZ09311 Human mac
C	40	13	59.1	23	14	ADV78737	ADV78737 PCR prime
C	41	13	59.1	24	2	AAV58821	AAV58821 Multimeri
C	42	13	59.1	24	6	ABZ21805	ABZ21805 PCR prime
C	43	13	59.1	25	13	ADR00162	ADR00162 DAPIK1 pro
C	44	13	59.1	26	4	ABL41661	ABL41661 Primer #2
C	45	13	59.1	26	13	ADV69164	ADV69164 Human per

ALIGNMENTS

RESULT 1	ADT97907	standard; DNA; 22 BP.
ID	ADT97907	
XX	ADT97907;	
AC		
XX		
DT	27-JAN-2005	(first entry)
XX		
DE	Human FGF-3 promoter 5' proximal region C allele probe.	
XX		
KW	Human; 88; fibroblast growth factor-3; FGF-3; cancer; oesophageal cancer;	
KW	breast cancer; ovarian cancer; prostate cancer; head and neck cancer;	
KW	oesophageal squamous cell carcinoma; PCR; probe; SNP;	
KW	single nucleotide polymorphism.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	modified_base 1	/*tag= a
FT	/*tag= a	/mod_base= OTHER
FT	/note= "C is covalently linked to a 6-carbo fluorescein moiety (FAM)"	
FT	modified_base 22	/*tag= b
FT	/*tag= b	/mod_base= OTHER
FT	/note= "G is covalently linked to a TAMRA moiety (6-carboxyretremethyl-rhodamine)"	
PN	US2004219582-A1.	
XX		
PD	04-NOV-2004.	
XX		
PF	11-MAR-2004; 2004US-00798652.	
XX		
PR	17-MAR-2003; 2003US-0455689P.	
XX		
PA	(GUOY/)	
XX		
PI	Guo Y;	
XX		
DR	WPI; 2004-794435/78.	

XX Novel isolated nucleic acid molecule having single nucleotide
PT polymorphism in upstream untranslated region of fibroblast growth factor-
PT 3 gene, useful for assessing related susceptibility of mammal to cancer.
PS
XX Claim 20; SEQ ID NO 7; 15bp; English.

XX The invention relates to an isolated nucleic acid molecule comprising the
CC upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3
CC gene sequence (ADT97901) with a single nucleotide polymorphism (SNP) at
CC position 69. Also included are an isolated nucleic acid molecule
CC comprising a sequence complementary to ADT97901, a vector comprising the
CC ADT97901 operably linked to a reporter gene, a host cell comprising the
CC vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating
CC a nucleic acid sample from the mammal, and determining whether a cytosine
CC or thymine is present at position 69 of ADT97901), a kit for performing
CC the method (comprising a first oligonucleotide probe which anneals
CC specifically with the target portion of the mammal's genome, where the
CC first probe comprises a first fluorescent label and a first fluorescence
CC quencher attached to its separate nucleotide residues and the target
CC portion includes the nucleotide residue located at position 69 of
CC ADT97901, and a pair of primers for amplifying a reference portion of the
CC FGF-3 gene, where the reference portion includes the nucleotide residue
CC located at position 69 of ADT97901) and a microarray having at least one
CC oligonucleotide probe that can anneal with a target portion of a mammal's
CC genome, where the target portion includes the nucleotide residue located
CC at position 69 of ADT97901. The method is useful for detecting SNP in FGF
CC -3 gene in a mammal, preferably a human, and is also useful for assessing
CC the relative susceptibility of a mammal to cancer (especially oesophageal
CC cancer), which shows an association with the presence of the C-allele.
CC The cancer is chosen from oesophageal, breast, ovarian, prostate, head
CC and neck cancer. The oesophageal cancer is oesophageal squamous cell
CC carcinoma. The present sequence is a Taqman PCR probe used to genotype
CC individuals for the C SNP in the UTR of FGF-3.
XX
SQ Sequence 22 BP; 3 A; 12 C; 3 G; 4 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 22; Length 22;
XX Best Local Similarity 100.0%; Pred. No. 13;
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTCCTCGCTCCAGCCACATG 22
DB |||||
1 CTCCTCGCTCCAGCCACATG 22

RESULT 2
ADT97906
ID ADT97906 standard; DNA; 22 BP.
XX
AC ADT97906;
XX
DT 27-JAN-2005 (first entry)
XX
XX Human FGF-3 promoter 5' proximal region T allele probe.
DE
XX Human; ss: fibroblast growth factor-3; FGF-3; cancer; oesophageal cancer;
KW breast cancer; ovarian cancer; prostate cancer; head and neck cancer;
KW oesophageal squamous cell carcinoma; PCR; probe; SNP;
KW single nucleotide polymorphism.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT 1
FT modified_base 1
FT /tag= a
FT /mod_base= OTHER
FT /note= "modified with Vic (not defined)"
FT
FT modified_base 22
FT /*tag= b
FT /mod_base= OTHER
FT /note= "G is covalently linked to a TAMRA moiety (6-
FT carbotetramethyl-rhodamine)"
OS

XX US2004219582-A1.
PN
XX 04-NOV-2004.
PD
XX
XX 11-MAR-2004; 2004US-00798652.
PF
XX
XX 17-MAR-2003; 2003US-0455689P.
PR
XX
XX (GUOY/) GUO Y.
PA
XX
XX Guo Y;
PI
XX WPI; 2004-794435/78.
DR
XX
XX Novel isolated nucleic acid molecule having single nucleotide
PT polymorphism in upstream untranslated region of fibroblast growth factor-
PT 3 gene, useful for assessing related susceptibility of mammal to cancer.
PS
XX Claim 34; SEQ ID NO 6; 15bp; English.

XX The invention relates to an isolated nucleic acid molecule comprising the
CC upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3
CC gene sequence (ADT97901) with a single nucleotide polymorphism (SNP) at
CC position 69. Also included are an isolated nucleic acid molecule
CC comprising a sequence complementary to ADT97901, a vector comprising the
CC ADT97901 operably linked to a reporter gene, a host cell comprising the
CC vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating
CC a nucleic acid sample from the mammal, and determining whether a cytosine
CC or thymine is present at position 69 of ADT97901), a kit for performing
CC the method (comprising a first oligonucleotide probe which anneals
CC specifically with the target portion of the mammal's genome, where the
CC first probe comprises a first fluorescent label and a first fluorescence
CC quencher attached to its separate nucleotide residues and the target
CC portion includes the nucleotide residue located at position 69 of
CC ADT97901, and a pair of primers for amplifying a reference portion of the
CC FGF-3 gene, where the reference portion includes the nucleotide residue
CC located at position 69 of ADT97901) and a microarray having at least one
CC oligonucleotide probe that can anneal with a target portion of a mammal's
CC genome, where the target portion includes the nucleotide residue located
CC at position 69 of ADT97901. The method is useful for detecting SNP in FGF
CC -3 gene in a mammal, preferably a human, and is also useful for assessing
CC the relative susceptibility of a mammal to cancer (especially oesophageal
CC cancer), which shows an association with the presence of the C-allele.
CC The cancer is chosen from oesophageal, breast, ovarian, prostate, head
CC and neck cancer. The oesophageal cancer is oesophageal squamous cell
CC carcinoma. The present sequence is a Taqman PCR probe used to genotype
CC individuals for the T SNP in the UTR of FGF-3.
XX
SQ Sequence 22 BP; 4 A; 12 C; 2 G; 4 T; 0 U; 0 Other;

XX Query Match 92.7%; Score 20.4; DB 13; Length 22;
XX Best Local Similarity 95.5%; Pred. No. 60;
XX Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CTCCTCGCTCCAGCCACATG 22
DB |||||
1 CTCCTCGCTCCAGCCACATG 22

RESULT 3
AEA51146
ID AEA51146 standard; DNA; 20 BP.
XX
XX AEA51146;
AC
XX
XX 11-AUG-2005 (first entry)
DT
XX
XX pyranair primer, SEQ ID NO: 20 to construct pyrF selection marker system.
DE
XX Expression; protein production; selectable marker; ss; pyrF; PCR; primer.
KW
XX Unidentified.
OS

XX WO2005052151-A1.
 PN 19-NOV-2003; 2003US-0523420P.
 XX 16-JAN-2004; 2004US-0537147P.
 PD 09-JUN-2005.
 XX 19-NOV-2004; 2004WO-US038884.
 PF 19-NOV-2003; 2003US-0523420P.
 XX 16-JAN-2004; 2004US-0537147P.
 XX (DOMC) DOM GLOBAL TECHNOLOGIES INC.
 PA Schneider JC, Chew LC, Badgley AK, Ramseier TM;
 XX MPI, 2005-417995/42.
 DR
 XX
 PT New auxotrophic Pseudomonas cell for use in a bacterial expression system
 PT that comprises a nucleic acid construct encoding a polypeptide that
 PT restores prototrophy to the host cell, useful for producing of
 PT recombinant polypeptides.
 CC
 XX
 PS Example 1; SEQ ID NO 20; 121pp; English.
 CC The present invention provides an improved expression system for the
 CC production of recombinant polypeptides utilizing auxotrophic Pseudomonas
 CC cell selectable markers. The invention also provides improved recombinant
 CC protein production in host cells through improved regulation of
 CC expression. The present sequence is the PCR primer used to construct pyrF
 CC selection marker system in a Pseudomonas fluorescens host cell expression
 CC system.
 CC
 XX
 SQ Sequence 20 BP; 4 A; 10 C; 2 G; 4 T; 0 U; 0 Other;
 Query Match 70.0%; Score 15.4; DB 14; Length 20;
 Best Local Similarity 94.1%; Pred. No. 7e+03; 1; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 5 CTCGCCTCCAGCAGCAT 21
 Db 1 CTCGCCTCCAGCAGCAT 17
 RESULT 4
 ADM79585
 ID ADM79585 standard; DNA; 20 BP.
 AC ADM79585;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Human ARTS1 related G446A mutation detection primer SEQ ID NO:14.
 XX
 KW human; ARTS1; ADP-riboosylation factor-like tumour suppressor 1;
 KW tumour suppressor; chromosome 13; cytosolic;
 KW abnormal cell proliferation inhibitor; caspase-1 protease modulator;
 KW mutation; detection; PCR; primer; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 PN WO2004033659-A2.
 PD 22-APR-2004.
 XX
 PF 10-OCT-2003; 2003WO-US032270.
 XX
 PR 11-OCT-2002; 2002US-0417842P.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX Croce CM;
 PI
 XX MPI; 2004-330451/30.
 DR

XX Novel ADP-riboosylation factor-like tumor suppressor 1 (ARTS1) protein
 PT useful for identifying modulators of Caspase-1 protease protein activity
 PT and for preventing abnormal cell growth in mammalian subjects.
 PT
 XX
 PS Example; SEQ ID NO 14; 47pp; English.
 CC
 XX The present invention describes the human ARTS1 (ADP-riboosylation factor-
 CC like tumour suppressor 1) protein (I). The human ARTS1 gene is located on
 CC chromosome 13, more specifically to 13q14. Also described: (1) an
 CC isolated nucleic acid molecule (II) comprising a sequence that encodes
 CC (I); (2) an isolated nucleic acid molecule (III) comprising the 3791
 CC nucleotide sequence of SEQ ID NO:2, or its fragment having at least 10
 CC nucleotides; (3) a recombinant expression vector (IV) comprising (III);
 CC (4) a host cell comprising (IV); (5) an oligonucleotide molecule (V)
 CC comprising a nucleotide sequence complementary to a nucleotide sequence
 CC of at least 5 nucleotides of SEQ ID NO:2; and (6) an isolated antibody
 CC (VI) that binds to an epitope on SEQ ID NO:1. Human ARTS1 has cytosolic
 CC activity, and can be used as an inhibitor of proliferation of abnormal
 CC cells. (I) is useful for identifying modulators of caspase-1 protease
 CC protein activity which involves performing a test assay by contacting a
 CC caspase-1 protease protein with a caspase-1 substrate such as FXBP46
 CC protein in the presence of a test compound, determining the level of
 CC processing of the substrate by the protease, and comparing the level to
 CC the level of processing of a caspase-1 substrate by caspase-1 protease
 CC protein in the absence of the test compound. (I) and (II) are useful for
 CC preventing abnormal cell growth in mammalian subjects. (II) is useful for
 CC designing probes and primers that are useful for detecting mutants of
 CC ARTS1 protein. (V) is useful as a probe in oligonucleotide hybridisation
 CC methods for identifying ARTS1. The present sequence represents a PCR
 CC primer for the detection of the G446A mutation in human ARTS, which is
 CC used in an example from the present invention.
 CC
 XX
 SQ Sequence 20 BP; 4 A; 8 C; 4 G; 4 T; 0 U; 0 Other;
 Query Match 67.3%; Score 14.8; DB 12; Length 20;
 Best Local Similarity 88.9%; Pred. No. 1.2e+04;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 5 CTCGCCTCCAGCAGCATG 22
 Db 2 CTCGCCTCCAGCAGCATG 19
 RESULT 5
 ADY00944/c
 ID ADY00944 standard; DNA; 29 BP.
 AC ADY00944;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE PCR primer 170 used to amplify human GALE region SNP DNA.
 XX
 KW SNP detection; breast tumor; endocrine disease;
 KW gynecology; and obstetrics; neoplasia; cytosolic; metastasis;
 KW gene therapy; RNA interference; UDP-galactose 4-epimerase; GALE;
 KW UDP-glucose 4-epimerase; lysophospholipase; HT014; LOC148902; LYPLA2; ss;
 KW PCR; primer.
 XX
 OS Homo sapiens.
 OS
 PN WO2005014846-A2.
 PD 17-FEB-2005.
 XX
 PF 27-MAY-2004; 2004WO-US016939.
 XX
 PR 24-JUL-2003; 2003US-0490234P.
 XX 25-NOV-2003; 2003US-00723681.
 PR 25-NOV-2003; 2003US-0525239P.
 XX
 PA (SEQU-) SEQUENOM INC.
 DR

```

XX  Roch RB, Nelson MR, Braun A, Kammerer SM, Reneland R;
P1  Hoyal-Wrightson CR;
XX
DR  WPI, 2005-163257/17.
XX
PT  Identifying risk of, preventing and/or treating breast cancer by
PT  identifying and/or analyzing polymorphic variations in nucleotide
PT  sequences within the human genome.
XX
PS  Example 15; Page 208; 617pp; English.
XX
CC  The invention relates to a novel method for identifying a subject at risk
CC  of breast cancer comprising detecting the presence or absence of a
CC  polymorphic variation associated with breast cancer. The method of the
CC  invention demonstrates cytostatic activity and may be useful for
CC  identifying a risk of, preventing and/or treating breast cancer and
CC  cancer metastasis. The methods may be utilized for gene therapy or RNA
CC  interference. The current sequence is that of a PCR primer of the
CC  invention which was used to amplify a human HT014/UC148902/LYP1A2
CC  (lysophospholipase II)/GALB (galactose-4-epimerase, UDP-) DNA containing
XX  a single nucleotide polymorphism (SNP).
SQ  Sequence 29 BP; 8 A; 1 C; 16 G; 4 T; 0 U; 0 Other;

Query Match          66.4%; Score 14.6; DB 14; Length 29;
Best Local Similarity 81.0%; Pred. No. 1.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0.

Oy      1 CTCCTCGCTCCAGCCACAT 21
        ||| ||| ||| ||| |||
Db      28 CTCCTCTCTCTCCCTCCACAT 8

RESULT 6
AAT90350
ID      AAT90350 standard; DNA, 30 BP.
XX
AC      AAT90350;
XX
DT      19-JAN-1998 (first entry)
XX
DE      DNA encoding peptide immunoreactive with HuMAb L94.
XX
KW      Tumour associated antigen; melanoma; cancer; therapy; immunotherapy;
XX      diagnosis; prognosis; vaccine; ss.
XX
OS      Homo sapiens.
XX
PN      WO9714795-A1.
XX
PD      24-APR-1997.
XX
PF      19-OCT-1995; 95WO-US013331.
XX
PR      19-OCT-1995; 95WO-US013331.
XX
PA      (WAYN-) WAYNE CANCER INST JOHN.
XX
PI      Irie RF, Kikumoto Y;
XX
DR  WPI, 1997-245112/22.
DR  P-PsDB; AAW26567.
XX
PT  Antibody against polypeptide having C-terminal amino acid sequence
PT  comprising Ala-Pro, Gly-Pro or Pro-Pro - useful to purify polypeptide
PT  which can be useful in human melanoma treatment.
XX
PS  Example 1; Page 40; 87pp; English.
XX
CC  This nucleotide sequence codes for a peptide (see AAW26567) that is
CC  immunoreactive with human monoclonal antibody WCI L94 that reacts to a
CC  human melanoma cell line. It was obtained by screening melanoma M14 and

```

CC	M12 expression libraries with L94. Isolated clones (see AAT90347-51)
CC	encode peptides (AA626566-73) that share the C-terminal sequence Ala-Pro.
CC	These peptides can be in an antigen composition to elicit an immune
CC	response, preferably a cytotoxic T lymphocyte response, specifically as a
CC	polyvalent tumour cell vaccine in conjunction with a cancer whole cell
CC	vaccine therapy for human melanoma. The peptides can also be used to
CC	enhance an immune response, treat and diagnose cancer
SQ	Sequence 30 BP; 5 A; 12 C; 8 G; 5 T; 0 U; 0 Other;
XX	
Query Match	64.5%; Score 14.2; DB 2; Length 30;
Best Local Similarity	84.2%; Pred. No. 2.2e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
OY	4 CCTGCCTTCAGCCACATG 22 DB 11 CGCGCCTTCAGCTCATG 29
RESULT 7	
ADB67764/C	
ID	ADB67764 standard; DNA; 25 BP.
XX	
AC	ADB67764;
XX	
DT	04-DEC-2003 (first entry)
XX	
DE	HIV-1 gp120 optimised sequencing primer g120c-SF4.
XX	
KM	HIV, Tct; HIV-1; Nef; gp120; vaccine; immunisation; human; anti-HIV;
XX	primer; ss.
KW	
OS	Synthetic.
XX	
XX	Human immunodeficiency virus 1.
XX	
FN	WO2003011334-A1.
PD	
XX	
PD	13-FEB-2003.
PF	26-JUL-2002; 2002WO-EP008343.
XX	
PR	27-JUL-2001; 2001GB-00018367.
XX	
PA	(GLAX) GLAXOSMITHKLINE BIOLOGICALS SA.
PA	(GLAX) GLAXO GROUP LTD.
PI	Erl Ff, Tite Jf, Van Wely CA, Voos G;
DR	WI; 2003-239474/23.
XX	
PT	
PT	Use of an HIV Tat, Nef or gp120 protein or polynucleotide in
PT	manufacturing a vaccine for a prime-boost delivery for the prophylactic
PT	or therapeutic immunization of humans against HIV.
XX	
XX	Example 15; Fig 21; 108bp; English.
XX	
XX	The present invention describes the use of an HIV Tat protein or
XX	polynucleotide, an HIV Nef protein or polynucleotide, or an HIV Tat
XX	protein or polynucleotide linked to an HIV Nef protein or polynucleotide,
XX	and an HIV gp120 protein or polynucleotide in manufacturing a vaccine for
XX	a prime-boost delivery for the prophylactic or therapeutic immunisation
XX	of humans against HIV. The protein or polynucleotide is delivered via a
XX	bombardment approach. Also described: (1) a recombinant DNA molecule
XX	comprising a Nef and/or Tat and/or gp120 gene in a vector in which the
XX	gene of interest is inserted 3' to an enhanced HCMV IE1 promoter; (2)
XX	particles, preferably gold particles, coated with recombinant DNA
XX	comprising at least two different vaccine compositions including: (a) a kit
XX	comprising at least two different vaccine compositions including: (a) a kit
XX	composition comprising particles coated with DNA encoding gp120 and nef
XX	and/or tat or nef/tat, and (b) a composition comprising gp120 and nef
XX	and/or tat or nef-tat DNA or proteins, where the DNA or proteins are not
XX	coated onto the particles. The HIV proteins or polynucleotides are used
XX	to produce vaccines having anti-HIV activity. The recombinant DNA
XX	

CC molecule can be used in manufacturing an HIV vaccine for the prophylactic
CC or therapeutic immunisation of humans. The present sequence represents an
CC oligonucleotide primer for sequencing optimised gp120, which is used in
CC an example from the present invention.

SO Sequence 25 BP; 6 A; 4 C; 12 G; 3 T; 0 U; 0 Other;

Query Match 63.6%; Score 14; DB 10; Length 25;
Best Local Similarity 77.3%; Pred. No. 2.7e+04;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CTCCTCGCTCCAGCCACATG 22
DB 22 CTTCCCACTCTCTGCACATG 1

RESULT 8

AEAS1123
ID AEAS1123 standard; DNA; 20 BP.

AC AEAS1123;

XX 11-AUG-2005 (first entry)

DE Human class II BHLH protein SCX gene specific sense primer.

XX Cell therapy; stem cell; trauma; tranquilizer or vulnerary; injury;
KW periodontal disease; antiinflammatory; mouth disease; scleraxis gene;
primer; ss.

OS Homo sapiens.

PN WO2005052140-A2.

XX 09-JUN-2005.

XX 22-NOV-2004; 2004WO-US039248.

XX 20-NOV-2003; 2003US-052602P.

XX (USSH) US DEPT OF HEALTH.

XX Shi S, Seo B, Miura M;

XX WPI; 2005-405376/41.

PT Novel isolated periodontal ligament multipotent stem cell, useful for
PT generating periodontal tissue or cells in a subject to reduce or
PT ameliorate trauma such as periodontal disease or physical injury due to
PT dental procedure.

PS Example 1; SEQ ID NO 7; 62pp; English.

XX The present invention relates to postnatal periodontal ligament stem
CC cells (PDLSCs). The invention is useful for periodontal tissue
CC regeneration, periodontal disease treatment and differential use of the
CC cells and methods of tissue cryopreservation. The invention is also
CC useful in cell therapy and in medical therapy e.g. for treating trauma.
CC The present sequence is the human class II BHLH protein scleraxis (SCX)
CC gene specific primer. This sequence is used in the isolation and
CC characterization of PDLSCs.

XX Sequence 20 BP; 3 A; 9 C; 3 G; 5 T; 0 U; 0 Other;

Query Match 62.7%; Score 13.8; DB 14; Length 20;
Best Local Similarity 88.2%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 CTCGCTCCAGCCACAT 21
DB 1 CTGGCTCCAGCTACAT 17

RESULT 9

ABO00949/C
ID ABO00949 standard; DNA; 24 BP.

XX ABO00949;

XX 11-JUN-2002 (first entry)

DE Oligonucleotide adapter/capture probe 940.

XX Oligonucleotide array; adapter sequence; probe; ss.

XX Synthetic.

PN WO200216649-A2.

XX 28-FEB-2002.

XX 27-AUG-2001; 2001WO-US026519.

XX 25-AUG-2000; 2000US-0227948P.

XX 29-AUG-2000; 2000US-0228854P.

XX (ILLU-) ILLUMINA INC.

XX Gunderson K;

XX WPI; 2002-292068/33.

PT Array comprising adapter sequences useful for immobilizing or detecting a
PT target nucleic acid sequence, has different addresses comprising
PT different specific capture probes.

PS Claim 1; Page 66; 261pp; English.

XX The invention relates to an oligonucleotide array (I) comprising at least
CC 25 different addresses (adapter sequences) with each comprising a
CC different capture probe selected from a group consisting of the sequences
CC given in ABO00010-ABQ13409. (I) is useful for immobilising a target
CC nucleic acid sequence by attaching a adapter nucleic acid (ABQ0010-
CC ABQ13409) to a target nucleic acid to form a modified target nucleic acid
CC and contacting the modified target nucleic acid with (II). The steps of
CC above method is useful for detecting a target nucleic acid, which further
CC comprises detecting the presence of the modified target nucleic acid

XX Sequence 24 BP; 5 A; 4 C; 8 G; 7 T; 0 U; 0 Other;

Query Match 62.7%; Score 13.8; DB 6; Length 24;
Best Local Similarity 88.2%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 CTCGCTCCAGCCACAT 21
DB 21 CTCGCTCCAGCCACAT 5

RESULT 10

ABQ05868
ID ABQ05868 standard; DNA; 24 BP.

XX ABQ05868;

XX 11-JUN-2002 (first entry)

DE Oligonucleotide adapter/capture probe 5859.

XX Oligonucleotide array; adapter sequence; probe; ss.

XX Synthetic.

XX WO200216649-A2.

XX 28-FEB-2002.

XX 27-AUG-2001; 2001WO-US026519.
 XX
 XX 25-AUG-2000; 2000US-0227948P.
 PR 29-AUG-2000; 2000US-0228854P.
 XX
 PA (ILLU-) ILLUMINA INC.
 XX
 PI Gunderson K;
 XX
 DR WPI; 2002-292068/33.
 XX
 PT Array comprising adapter sequences useful for immobilizing or detecting a
 PT target nucleic acid sequence, has different addresses comprising
 PT different specific capture probes.
 XX
 PS Claim 1, Page 161, 261pp; English.

XX The invention relates to an oligonucleotide array (I) comprising at least
 CC 25 different addresses (adapter sequences) with each comprising a
 CC different capture probe selected from a group consisting of the sequences
 CC given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
 CC nucleic acid sequence by attaching a adapter nucleic acid (ABQ00010-
 CC ABQ13409) to a target nucleic acid to form a modified target nucleic acid
 CC and contacting the modified target nucleic acid with (I). The steps of
 CC above method is useful for detecting a target nucleic acid, which further
 CC comprises detecting the presence of the modified target nucleic acid
 XX

SO Sequence 24 BP; 7 A; 8 C; 4 G; 5 T; 0 U; 0 Other;
 Query Match 62.7%; Score 13.8; DB 6; Length 24;
 Best Local Similarity 88.2%; Pred. No. 3.2e+04;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CTCGCCTCCAGCACAT 21
 |||||
 DB 4 CTCGCCTCCAGCACAT 20

RESULT 11
 ABQ05827/c
 ID ABQ05827 standard; DNA; 24 BP.
 XX
 AC ABQ05827;
 XX
 DT 11-JUN-2002 (first entry)
 XX
 DE Oligonucleotide adapter/capture probe 5818.
 XX
 KW Oligonucleotide array; adapter sequence; probe; ss.
 XX
 OS Synthetic.
 XX
 PN WO200216649-A2.
 PD 28-FEB-2002.
 XX
 PF 27-AUG-2001; 2001WO-US026519.
 XX
 PR 25-AUG-2000; 2000US-0227948P.
 PR 29-AUG-2000; 2000US-0228854P.
 XX
 PA (ILLU-) ILLUMINA INC.
 XX
 PI Gunderson K;
 XX
 DR WPI; 2002-292068/33.
 XX
 PT Array comprising adapter sequences useful for immobilizing or detecting a
 PT target nucleic acid sequence, has different addresses comprising
 PT different specific capture probes.
 XX
 PS Claim 1, Page 161, 261pp; English.

XX The invention relates to an oligonucleotide array (I) comprising at least
 CC 25 different addresses (adapter sequences) with each comprising a
 CC different capture probe selected from a group consisting of the sequences
 CC given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
 CC nucleic acid sequence by attaching a adapter nucleic acid (ABQ00010-
 CC ABQ13409) to a target nucleic acid to form a modified target nucleic acid
 CC and contacting the modified target nucleic acid with (I). The steps of
 CC above method is useful for detecting a target nucleic acid, which further
 CC comprises detecting the presence of the modified target nucleic acid
 XX

SO Sequence 24 BP; 5 A; 4 C; 8 G; 7 T; 0 U; 0 Other;
 Query Match 62.7%; Score 13.8; DB 6; Length 24;
 Best Local Similarity 88.2%; Pred. No. 3.2e+04;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CTCGCCTCCAGCACAT 21
 |||||
 DB 21 CTCGCCTCCAGCACAT 5

RESULT 12
 AAS07411
 ID AAS07411 standard; DNA; 20 BP.
 XX
 AC AAS07411;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE PCR primer K21L19L used to map the REVOLUTA gene.
 XX
 KW Revoluta; Rev; corn; barley; rice; tomato; PCR primer; apical meristem;
 KW leaf; floral organ; stem; transgenic plant; crop yield; cereal; fruit;
 KW pharmaceutical; industrial; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200133944-A1.
 PD 17-MAY-2001.
 XX
 PF 10-NOV-2000; 2000WO-US030794.
 XX
 PR 10-NOV-1999; 99US-0164587P.
 XX
 PA (SLAD/) SLADE A.
 PA (MADI/) MADISEN L.
 PA (COMA/) COMAI L.
 XX
 PI Slade A, Madisen L, Comai L;
 XX
 DR WPI; 2001-328861/34.
 XX
 PT Isolated DNA molecule comprising a sequence that encodes a REVOLUTA
 PT protein, useful for producing transgenic plants with modulated cell
 PT division.
 XX
 PS Example 1; Page 48, 149pp; English.

XX AAS07401-AAS07571 represent REVOLUTA (REV) coding sequences and PCR
 CC primers of the invention. The REV nucleic acid sequences were isolated
 CC from plants such as Arabidopsis thaliana, tomato, corn, barley and rice.
 CC The REV gene is required to promote the growth of apical meristems, but
 CC has an opposite effect on meristems of leaves, floral organs and stems,
 CC such that it acts to limit cell division reducing the rate of plant
 CC growth and final size of the tissue. Therefore, loss of functional REV
 CC leads to increases in the size of floral organs, leaf and stem tissue.
 CC DNA encoding the REV protein is useful for modulating plant cell
 CC division. The mutant REV DNA is also useful for producing transgenic
 CC plants with modulated cell division. These transgenic plants can be used
 CC to increase crop yield in cereals and fruits, and as a potential source
 CC of pharmaceuticals and industrial products

XX Sequence 20 BP; 4 A; 10 C; 1 G; 5 T; 0 U; 0 Other;
SQ

Query Match 61.8%; Score 13.6; DB 4; Length 20;
Best Local Similarity 80.0%; Pred. No. 3.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACA 20
Db 1 CTCCTCTTTCACAGACACA 20

RESULT 13
AAX57861
ID AAX57861 standard; DNA; 22 BP.
XX
AC AAX57861;
XX
DT 15-JUL-1999 (first entry)
XX
DE PCR primer used in construction of yeast artificial chromosome.
XX
KM YAC; yeast artificial chromosome; PCR primer; sexual dysfunction;
KM reporter gene; transgenic mammal; therapy; circadian function;
KM sleep disorder; eating disorder; premenstrual syndrome; birth defect;
KM autoimmune disorder; ss.
XX
OS Synthetic.
XX
PN GB2331752-A.
XX
PD 02-JUN-1999.
XX
PF 27-NOV-1998; 98GB-00026126.
XX
PR 28-NOV-1997; 97GB-00025311.
PR 28-NOV-1997; 97GB-00025313.
PR 20-MAR-1998; 98GB-00006072.
PR 05-NOV-1998; 98GB-00024275.
XX
PA (MED1-) MEDICAL RES COUNCIL.
XX
PI Shen S, Schedl A, Harmar AJ;
XX
DR WPI; 1999-290603/25.
XX
PT New reporter gene labeled YAC vectors and transgenic mammals used for
PT screening potential active agents.
XX
PS Disclosure; Page 56; 98pp; English.
XX
CC This sequence represents a PCR primer used in the construction of a yeast
CC artificial chromosome of the invention. The yeast artificial chromosome
CC (YAC) vectors contain a reporter gene and transgenic mammals produced
CC using them may be used to screen for an agent affecting nucleotide
CC expression and gives easier monitoring of in vivo expression. The vector
CC is used in the production of transgenic mammals for testing potential
CC pharmaceutical or veterinary agents. pYAM4 is used to amplify YAC. The
CC assay may be used to screen for agents useful in treatment of disturbance
CC of circadian function, sleep disorders, eating disorders, premenstrual
CC syndrome, autoimmune disorders, birth defects in women and/or sexual
CC dysfunction. The agents thus detected may be used for treatment of
CC disorders related to the expression pattern of a nucleotide such as those
CC above. The vectors have more concentrated YAC DNA, which allows better
CC and more reliable gene transfer. The presence of a reporter gene allows
CC easy monitoring of in vivo expression and the vectors allow for gene
CC overexpression (3-5 fold) and easy site determination. The pYAM4
CC amplification vector does not contain the thymidine kinase gene, which
CC causes male infertility in transgenic mice
XX
SQ Sequence 22 BP; 5 A; 11 C; 0 G; 6 T; 0 U; 0 Other;
Query Match 61.8%; Score 13.6; DB 2; Length 22;

Best Local Similarity 80.0%; Pred. No. 3.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TCCTCGCCTCCAGCCACAT 21
Db 1 TTCATCACCTCCATCCACAT 20

RESULT 14
AAC85242
ID AAC85242 standard; DNA; 22 BP.
XX
AC AAC85242;
XX
DT 22-MAR-2001 (first entry)
XX
DE Reverse primer D 26376 for determining size of YAC insert.
XX
KM Internal ribosomal entry site; IRES; yeast artificial chromosome; YAC;
KM vector; centromere; telomere; origin of replication; transgenic;
KM circadian function; sleep disorder; eating disorder;
KM premenstrual syndrome; autoimmune disease; birth defect;
KM sexual dysfunction; serotonin transporter; VIP2 receptor; SERT; VIPR2;
KM polymerase chain reaction; PCR; primer; amplify; YAC 35D8/D6;
KM YAC HSC7E526/V12; ss.
XX
OS Synthetic.
XX
PN GB2350613-A.
XX
PD 06-DEC-2000.
XX
PF 17-AUG-2000; 2000GB-00020335.
XX
PR 28-NOV-1997; 97GB-00025311.
PR 28-NOV-1997; 97GB-00025313.
PR 20-MAR-1998; 98GB-00006072.
PR 05-NOV-1998; 98GB-00024275*
PR 27-NOV-1998; 98GB-00026126.
XX
PA (MED1-) MEDICAL RES COUNCIL.
XX
PI Shen S, Schedl A, Harmar AJ;
XX
DR WPI; 2001-034098/05.
XX
PT Transgenic organism for identifying potential therapeutic agents able to
PT modulate gene expression, comprises a yeast artificial chromosome vector.
XX
PS Example; Page 54; 93pp; English.
XX
CC The sequences given in AAC85227-50 are primers which were used to
CC determine the size of the integrated YAC 35D8/D6 and YAC HSC7E526/V12
CC constructs in a transgenic founder animal. The constructs were prepared
CC from novel yeast artificial chromosome (YAC) vectors each of which
CC comprises a centromere, two telomeres, at least one origin of
CC replication, an internal ribosomal entry site (IRES), and a selection
CC gene that is specifically removable from the vector. The resulting YAC
CC are used to produce transgenic organisms for use in screening for agents
CC that can affect the expression pattern of a nucleotide sequence of
CC interest (NOI) or the activity of its expression product. The identified
CC agents are potentially useful as pharmaceutical and veterinary agents for
CC treating disorders of circadian function; sleep or eating disorders;
CC premenstrual syndrome; autoimmune diseases; birth defects in women and/or
CC sexual dysfunction, and also as lead compounds for developing agents with
CC other activities. YAC can also be used for expression, regulation and/or
CC functional studies on NOI, in combination with other NOI, compounds or
CC compositions. The new vectors provide high YAC copy numbers and allow
CC easy monitoring (in vivo) of the expression pattern of NOI;
CC (over)expression of NOI and a reporter gene; and determination of the
CC sites where NOI is expressed. Incorporation of IRES allows expression of
CC at least two nucleic acid sequences (e.g. NOI plus a reporter of more
CC than one NOI)

```

XX SQ Sequence 22 BP; 5 A; 11 C; 0 G; 6 T; 0 U; 0 Other;
      Query Match          61.8%; Score 13.6; DB 5; Length 22;
      Best Local Similarity 80.0%; Pred. No. 3.9e+04;
      Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 TCCCTCGCCTCCAGCCACAT 21
    |||||
Db 1 TTCATCACCCTCCATCCACAT 20

```

```

      Query Match          61.8%; Score 13.6; DB 9; Length 25;
      Best Local Similarity 80.0%; Pred. No. 3.9e+04;
      Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 TCCCTCGCCTCCAGCCACAT 21
    |||||
Db 23 TCCCTCGCCTCCACCTACTT 4

```

Search completed: March 25, 2006, 19:02:19
Job time : 374.5 secs

```

RESULT 15
ACI88529/c
ID ACI88529 standard; DNA; 25 BP.
XX AC
XX ACI88529;
XX
XX 14-OCT-2003 (first entry)
XX
XX Human microarray DNA oligonucleotide SEQ ID NO 88520.
XX
XX EST; sg; probe; expressed sequence tag; microarray; gene expression;
XX genetic variation; biallelic marker; polymorphism; human;
XX cross-species comparison.
XX
XX Homo sapiens.
XX
XX US2003104410-A1.
XX
XX 05-JUN-2003.
XX
XX 15-MAR-2002; 2002US-00098263.
XX
XX 16-MAR-2001; 2001US-0276759P.
XX
XX (AFXY-) AFFYMETRIX INC.
XX
XX Miltmann MP;
XX
XX WPI; 2003-567953/53.
XX
XX New array of nucleic acid probes, useful for in situ hybridization, in
XX Southern, Northern or dot-blot hybridization to identify or detect the
XX sequence or specific mutations of any gene.
XX
XX Claim 1; SEQ ID NO 88520; 9pp; English.
XX
XX The invention discloses a microarray comprising a plurality of nucleic
XX acid probes including one of 2,018,500 fully defined sequences, or its
XX perfect match, perfect mismatch, antisense match or antisense mismatch.
XX Also disclosed is a method of gene expression analysis. The array is used
XX in monitoring gene expression levels by hybridisation to a DNA library,
XX in analysis of genetic variation or in hybridisation of tag-labelled
XX compounds. The nucleic acid probes are specifically designed for analysis
XX of at least one target sequence. The method of analysis comprises
XX hybridising at least one or more nucleic acids to at least two or more
XX nucleic acid probes and detecting the hybridisation. The nucleic acid
XX probes are attached to a solid support. The analysis comprises monitoring
XX gene expression levels, identifying biallelic markers or polymorphisms,
XX or family members of a gene and a cross-species comparison. Each of the
XX nucleic acids further comprises a tag sequence. The array of nucleic acid
XX probes is useful in situ hybridisation, in Southern, Northern or dot-
XX blot hybridisation to identify or detect the sequence or specific
XX mutations of any gene, in mapping the 5' termini of mRNA molecules by
XX primer extensions or in screening cDNA or genomic libraries or subclones
XX for additional subclones containing segments of DNA that have been
XX isolated and previously sequenced. The sequence presented is one of the
XX nucleic acid probes incorporated in the microarray. Note: The sequence
XX data for this patent can also be obtained in electronic format directly
XX from USPTO at seqdata.uspto.gov/sequence.html
XX
XX Sequence 25 BP; 9 A; 3 C; 10 G; 3 T; 0 U; 0 Other;
XX
XX

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 13:43:43 ; Search time 3041.5 Seconds
(without alignments)
338.424 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22
Sequence: 1 ctccctcgctccagcacacatg 22

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 52094

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_eac1:*
2: gb_eac2:*
3: gb_eac3:*
4: gb_hac:*
5: gb_eac4:*
6: gb_eac5:*
7: gb_eac6:*
8: gb_eac7:*
9: gb_gac1:*
10: gb_gac2:*
11: gb_gac3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14.2	64.5	25	10	CG730897 1119130F0
C 2	12.4	56.4	29	10	CG676355 PR10118b
C 3	12	54.5	24	9	AZ762096 1M055618
C 4	12	54.5	25	1	A1416870 sal1b03.x
C 5	12	54.5	25	8	D21049 HMG502034
C 6	12	54.5	29	9	AZ332711 1M0061M4
C 7	12	54.5	29	9	AZ788256 2M0035A17
C 8	12	54.5	29	10	C2477475
C 9	11.6	52.7	22	9	AZ980990 2M0258F17
C 10	11.4	51.8	24	9	AZ655652 1M0530D13
C 11	11.4	51.8	28	1	AA864650 ch37b09.s
C 12	11.4	51.8	28	1	A1158446
C 13	11.2	50.9	16	1	AM247040 2822414.5
C 14	11.2	50.9	22	9	AZ863707 2M0171M19
C 15	11.2	50.9	28	1	A1279523 q151d12.x
C 16	11.2	50.9	29	9	AZ498825 1M0336J14
C 17	11	50.0	22	2	BG926061 HNC23-1-E
C 18	11	50.0	22	9	AZ778032 2M0013A11
C 19	11	50.0	23	9	AZ785027 2M0028H03
C 20	11	50.0	24	9	AZ404465 1M0172P09
C 21	11	50.0	25	1	A1424019 cf51e01.x
C 22	11	50.0	27	9	AZ439789 1M0230004

C 23	11	50.0	27	9	BH851753
C 24	11	50.0	28	9	AZ937148
C 25	11	50.0	29	9	AZ666737
C 26	11	50.0	29	9	AZ780387
C 27	11	50.0	29	11	CT024781
C 28	11	50.0	30	2	BF099623
C 29	11	50.0	30	2	BE367782
C 30	11	50.0	30	8	DN955616
C 31	11	50.0	30	9	AZ634665
C 32	11	50.0	30	9	AZ864869
C 33	10.8	49.1	19	9	AZ485264
C 34	10.8	49.1	21	9	AZ348213
C 35	10.8	49.1	23	9	AZ627985
C 36	10.8	49.1	27	9	AZ797359
C 37	10.8	49.1	28	5	C00171
C 38	10.8	49.1	28	9	AZ343801
C 39	10.8	49.1	29	9	AZ608734
C 40	10.8	49.1	30	8	CV999848
C 41	10.8	49.1	30	9	AZ456295
C 42	10.8	49.1	30	10	C2194668
C 43	10.6	48.2	19	9	AZ381798
C 44	10.6	48.2	19	9	AZ782026
C 45	10.6	48.2	20	7	CF932153

ALIGNMENTS

RESULT 1
CG730897/c 25 bp DNA linear GSS 20-OCT-2003
LOCUS 1119130F05.2EL x1 1119 - Rescuem Grd AA Zea mays genomic, genomic
DEFINITION survey sequence.

ACCESSION CG730897
VERSION CG730897.1 GI:37773389
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
TITLE clade; Panicoideae; Andropogoneae; Zea.
JOURNAL 1 (bases 1 to 25)

COMMENT Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1119130 row: F column: 05
Class: transposon-tagged.

FEATURES
Location/Qualifiers
1..25
/organism="Zea mays"
/mol_type="genomic DNA"
/culturivar="mixed background W23/A18/B73/K55"
/db_xref="taxon:4577"
/feature_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="1119 - Rescuem Grd AA"
/note="Organ: leaf; Vector: Rescuem Grd AA"
PB1uescript backbone; Site 1: BamHI; Site 2: BglII;
Rescuem is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on Rescuem, go to the web
site 'www.zmbd.iastate.edu' and follow the links for

ACCESSION A1416870
VERSION A1416870.1 GI:4260374
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 25)
Shoemaker, R., Keim, P., Vocklin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. GENOME SYSTEMS
CLONE ID: Gm-cl005-102 Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand this clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: T7 ET from Amerisham
High quality sequence stop: 1
POLYA=No.

FEATURES
SOURCE Location/Qualifiers
1..25
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl005-102"
/lab_host="XL10-Gold"
/clone_1db="Gm-cl005"
/note="Vector: pBluescript II XE; Site 1: EcoRI; Site 2: XhoI. This library was constructed by Dr. Randy C. Shoemaker and Dr. John Erpelting, USDA-ARS Agronomy Department, G401 Agronomy Hall, Iowa State University, Ames, IA 50011-1010 Phone: 515-294-6233 FAX: 515-294-2299 email: rcsheo@iastate.edu"

ORIGIN
Query Match 54.5%; Score 12; DB 1; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.7e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
1 CTCCTGCGCTCCAGCCACA 20
6 CCCCCCCCCCTCCACCCCA 25

Db

RESULT 5
LOCUS D21049 25 bp mRNA linear EST 30-JUL-1996
DEFINITION HM06S02034 Human promyelocyte Homo sapiens cDNA clone mp0338 3', mRNA sequence.
D21049
D21049.1 GI:504869
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 25)
AUTHORS Okubo, K., Fukushima, A., Yoshii, J., Niyama, T., Kojima, Y., Yoshinari, H., Arimoto, J. and Matsubara, K.
TITLE Gene expression of human promyelocytic cell line HL60 before and after induction of differentiation. A new application of 3'directed cDNA sequencing
JOURNAL Unpublished (1993)
COMMENT Contact: Okubo, K., Fukushima, A., Yoshii, J., Niyama, T., Kojima, Y., Yoshinari, H., Arimoto, J. and Matsubara, K.
Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamada-oka, Suita, Osaka 565, Japan.
Location/Qualifiers
1..25
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="mp0338"
/clone_1db="Human promyelocyte"
/note="Female, adult, cell_line = HL60, cell_type = promyelocyte."

FEATURES
SOURCE

ORIGIN
Query Match 54.5%; Score 12; DB 8; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.7e+06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
1 CTCCTGCGCTCCAGCCACA 20
4 CGCCCCCAGCTCCAGCCTCA 23

Db

RESULT 6
LOCUS AZ332711/ 29 bp DNA linear GSS 29-SEP-2000
DEFINITION IM0061N14F Mouse 10kb plasmid UUGCIM library Mus musculus genomic clone UUGCIM0061N14 F, genomic survey sequence.
AZ332711
AZ332711.1 GI:10396624
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 29)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Irlam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Relliy, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D. Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0061 row: N column: 14
Seq primer: CGTTGTAAACGACGGCCACT
Class: plasmid ends
High quality sequence stop: 29.
Location/Qualifiers
1..29
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCIM0061N14"

Class: transposon insertion site.	location/Qualifiers	source
1.	.29	

ОБЪЕКТ

Query Match	54.5%	Score 12;	DB 10;	Length 29;
Best Local Similarity	75.0%	Pred. No. 1.7e+06;		
Matches 15; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

RESULT 9	
AZ980990/c	
LOCUS	AZ980990
DEFINITION	22 bp DNA linear GSS 27-APR-2001
ACCESSION	U00258F17 Mouse 10kb plasmid UUGC2M library Mus musculus genomic
VERSION	clone UUGC2M0258F17 F, genomic survey sequence.
KEYWORDS	AZ980990
SOURCE	AZ980990.1 GI:13852217
ORGANISM	GSS.
	Mus musculus (house mouse)
	Mus musculus

ORIGIN

Query Match	52.7%	Score 11.6;	DB 9;	Length 22;
Best Local Similarity	77.8%;	Pred. No. 2.4e+06;		
Matches 14; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0

RESULT	10
AZ655652/c	
LOCUS	
DEFINITION	24 bp DNA linear GSS 14-DEC-2000
ACCESSION	AZ655652
VERSION	I1M0530D1Jr Mouse 10kb plasmid tUGC1M library Mus musculus genomic
KEYWORDS	clone tUGC1M0530D1J R, genomic survey sequence.
SOURCE	AZ655652.1 GI:11792798
ORGANISM	GSS.
	Mus musculus (house mouse)
	Mus musculus

FEATURES	Location/Qualifiers
source	1. .22

FEATURES

SOURCE

RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

ORIGIN

Query Match 51.8%; Score 11.4; DB 1; Length 28;
Best Local Similarity 92.3%; Pred. No. 3e+06;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 7 CGCCTCCAGCCAC 19
|||||
Db 10 CGCCTACAGCCAC 22

RESULT 13

LOCUS AM247040 16 bp mRNA linear EST 07-JAN-2000
DEFINITION 2822414.Sprine NIH_MGC_7 Homo sapiens CDNA clone IMAGE:2822414 5',
mRNA sequence.

ACCESSION AM247040
VERSION AM247040.1 GI:6590033
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 16)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other ESTs: 2822414.3prine
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DTF CDNA Library Preparation: Ling
Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LIML) DNA Sequencing by: Berkeley MGC sequencing
project clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LIML at:
www.bio.liml.gov/btrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center.
Trimming: cross match from University of Washington Genome Center
PHRAP suite. Poly-T identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu Low Quality Sequence: 15
contiguous PHRED high quality bases following vector sequence. Very
Low Quality Sequence: Trace file contained 16 contiguous distinct
peaks following vector sequence.
Plate: LCM9 row: F column: 15
High quality sequence stop: 15.

FEATURES

Source

1..16
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2822414"
/issue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7, Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 50.9%; Score 11.2; DB 1; Length 16;
Best Local Similarity 81.2%; Pred. No. 3.6e+06;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 TCCTCGCCTCCAGCC 17
|||||
Db 16 TCCTTCCTCTGTC 1

RESULT 14

LOCUS AZ863707/c 22 bp DNA linear GSS 21-FEB-2001
DEFINITION 2M0171N19R Mouse 10kb plasmid UGCLM library Mus musculus genomic
clone UGSC2M0171N19 R, genomic survey sequence.

ACCESSION AZ863707
VERSION AZ863707.1 GI:13062279
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 22)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausen,A. and Wright,D., Weis,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weis
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0171 row: N column: 19
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 22.

FEATURES

Source

1..22
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGSC2M0171N19"
/sex="Male"
/lab_host="B. Coli strain XL10-Gold, T1-resistant, F-"
/clone_1lb="Mouse 10kb plasmid UGCLM library"
/note="Vector: pMD29v; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD29 (g14732114|gD|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E.coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 50.9%; Score 11.2; DB 9; Length 22;
Best Local Similarity 81.2%; Pred. No. 3.6e+06;

Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 CCTGCGCTCCAGCCA 18
 |||||
 19 CCATGCGCGCCACCA 4

RESULT 15
 A1279523 28 bp mRNA linear EST 29-JAN-1999
 LOCUS A1279523
 DEFINITION Q151d12.x1 Soares NHHMPu S1 Homo sapiens cDNA clone IMAGE:1875863
 3, similar to SW:EXTN_TOBAC P13983 EXTENSIN PRECURSOR ; contains
 MER22.b1 MSRI repetitive element ;, mRNA sequence.

ACCESSION A1279523
 VERSION A1279523
 KEYWORDS EST, GI:3917757
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE 1 (bases 1 to 28)
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Insert Length: 588 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.

FEATURES
 source location/Qualifiers

1..28
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1875863"
 /issue_type="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"
 /clone_lib="Soares NHHMPu S1"
 /note="Organ: mixed (see below); Vector: pT73D-Pac
 (Pharmacia) with a modified polylinker; Site_1: Not 1;
 Site_2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte ZNHM, pregnant uterus
 NHHMP, and fetal heart NHH19W) were mixed, and 88 circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from pools of
 5,000 clones made from the same 3 libraries. The pools
 consisted of I.M.A.G.E. clones 260232-265223,
 340488-345479, and 484488-489479."

ORIGIN

Query Match 50.9%; Score 11.2; DB 1; Length 28;
 Best Local Similarity 81.2%; Pred. No. 3.6e+06;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 CCTGCGCTCCAGCCAC 19
 |||||
 DB 12 CCTGCGCGCGCCAC 27

Search completed: March 25, 2006, 15:28:19
 Job time : 3043.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 12:58:33 ; Search time 129.5 Seconds
(without alignments)
301.980 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22
Sequence: 1 ctcctcgctccagccacacg 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 1026780

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:
1: /cgn2_6/pdata/1/ina/1-COMB.seq:*
2: /cgn2_6/pdata/1/ina/5-COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A-COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B-COMB.seq:*
5: /cgn2_6/pdata/1/ina/H-COMB.seq:*
6: /cgn2_6/pdata/1/ina/PCUS-COMB.seq:*
7: /cgn2_6/pdata/1/ina/PP-COMB.seq:*
8: /cgn2_6/pdata/1/ina/RB-COMB.seq:*
9: /cgn2_6/pdata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	13.6	61.8	27	2	US-08-460-890A-19 Sequence 19, Appl
C 2	13.6	61.8	27	2	US-08-460-890A-22 Sequence 22, Appl
C 3	13.6	61.8	27	3	US-08-167-641C-19 Sequence 19, Appl
C 4	13.6	61.8	27	3	US-08-167-641C-22 Sequence 22, Appl
C 5	13.6	61.8	27	3	US-08-460-971A-19 Sequence 19, Appl
C 6	13.6	61.8	27	3	US-08-460-971A-22 Sequence 22, Appl
C 7	13.6	61.8	27	3	US-08-462-040-19 Sequence 19, Appl
C 8	13.6	61.8	27	3	US-08-462-040-22 Sequence 22, Appl
C 9	13.4	60.9	25	3	US-09-396-196G-58648 Sequence 58648, A
C 10	13.4	60.9	25	3	US-09-396-196G-58649 Sequence 58649, A
C 11	13.4	60.9	25	3	US-09-396-196G-58650 Sequence 58650, A
C 12	13.4	59.1	23	2	US-08-766-982-10 Sequence 10, Appl
C 13	13.4	59.1	23	2	US-09-296-219-10 Sequence 10, Appl
C 14	13.4	59.1	24	2	US-08-529-190B-16 Sequence 16, Appl
C 15	13.4	59.1	30	3	US-09-119-507B-10 Sequence 10, Appl
C 16	13.4	59.1	30	3	US-09-119-507B-38 Sequence 38, Appl
C 17	13.4	59.1	30	3	US-08-897-556A-10 Sequence 10, Appl
C 18	13.4	59.1	30	3	US-08-897-556A-38 Sequence 38, Appl
C 19	13.4	59.1	30	3	US-09-547-693-10 Sequence 10, Appl
C 20	13.4	59.1	30	3	US-09-547-693-38 Sequence 38, Appl
C 21	12.8	58.2	17	3	US-08-944-410-88 Sequence 84, Appl
C 22	12.8	58.2	25	3	US-09-396-196G-98155 Sequence 98155, A
C 23	12.8	58.2	25	3	US-09-396-196G-98156 Sequence 98156, A
C 24	12.6	57.3	24	2	US-08-529-190B-14 Sequence 14, Appl

25	12.6	57.3	24	3	US-07-792-600-8 Sequence 8, Appl
26	12.6	57.3	24	3	US-09-157-021-8 Sequence 8, Appl
27	12.6	57.3	24	3	US-09-156-842-8 Sequence 8, Appl
28	12.6	57.3	24	3	US-09-591-514-8 Sequence 8, Appl
29	12.6	57.3	25	3	US-09-396-196G-2327 Sequence 2327, Ap
30	12.6	57.3	25	3	US-09-396-196G-44454 Sequence 44454, A
C 31	12.6	57.3	25	3	US-09-396-196G-76723 Sequence 76723, A
C 32	12.6	57.3	27	6	PCT-US93-02352-4 Sequence 4, Appl
C 33	12.6	57.3	29	6	PCT-US93-02352-6 Sequence 6, Appl
C 34	12.4	56.4	19	3	US-09-517-467B-9 Sequence 9, Appl
C 35	12.4	56.4	24	3	US-08-899-112B-17 Sequence 17, Appl
C 36	12.4	56.4	24	3	US-09-887-145-6 Sequence 6, Appl
C 37	12.4	56.4	24	3	US-09-011-553-18 Sequence 18, Appl
C 38	12.4	56.4	24	3	US-09-971-773-61 Sequence 61, Appl
C 39	12.4	56.4	25	2	US-08-268-020A-1 Sequence 1, Appl
C 40	12.4	56.4	25	3	US-08-899-112B-18 Sequence 18, Appl
C 41	12.4	56.4	25	3	US-09-011-553-19 Sequence 19, Appl
C 42	12.4	56.4	25	3	US-09-396-196G-43715 Sequence 43715, A
C 43	12.4	56.4	25	3	US-09-396-196G-43716 Sequence 43716, A
C 44	12.4	56.4	25	3	US-09-396-196G-56734 Sequence 56734, A
C 45	12.4	56.4	25	3	US-09-396-196G-61491 Sequence 61491, A

ALIGNMENTS

RESULT 1
US-08-460-890A-19/C
Sequence 19, Application US/08460890A
; Patent No. 5994109
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchaik, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESS: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: PARCISO for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,890A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Waiburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 488-1600
; TELEFAX: (213) 955-0440
; TELE: 67-3510
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:

LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-460-890A-19

Query Match 61.8%; Score 13.6; DB 2; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACA 20
DB 20 CTCCTCGCCTCCCTCCCA 1

RESULT 2
US-08-460-890A-22/c
Sequence 22, Application US/08460890A

PATENT No. 5994109
GENERAL INFORMATION:
APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,890A
CLASSIFICATION: 435
FILING DATE: June 5, 1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Wardburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-460-890A-22

Query Match 61.8%; Score 13.6; DB 2; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACA 20
DB 20 CTCCTCGCCTCCCTCCCA 1

RESULT 3
US-08-167-641C-19/c
Sequence 19, Application US/08167641C
Patent No. 6033884
GENERAL INFORMATION:

APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,641C
FILING DATE: December 14, 1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Wardburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 205/012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-167-641C-19

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACA 20
DB 20 CTCCTCGCCTCCCTCCCA 1

RESULT 4
US-08-167-641C-22/c
Sequence 22, Application US/08167641C
Patent No. 6033884
GENERAL INFORMATION:
APPLICANT: Woo, Savio L.C.

```

; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,641C
; FILING DATE: December 14, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 205/012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-167-641C-22

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACA 20
Db 20 CTCCTCGCCTCCAGCCACA 1

RESULT 5
US-08-460-971A-19/C
; Sequence 19, Application US/08460971A
; Patent No. 6150168
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
```

```

; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,971A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-460-971A-19

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACA 20
Db 20 CTCCTCGCCTCCAGCCACA 1

RESULT 6
US-08-460-971A-22/C
; Sequence 22, Application US/08460971A
; Patent No. 6150168
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,971A
```

FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-460-971A-22

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCTCCAGCCACA 20
DB 20 CTCCTCGCTCCAGCCACA 1

RESULT 7
US-08-462-040-19/c
Sequence 19, Application US/08462040
Patent No. 6177554
GENERAL INFORMATION:
APPLICANT: Moo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gotchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,040
FILING DATE: June 5, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-462-040-19

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCTCCAGCCACA 20
DB 20 CTCCTCGCTCCAGCCACA 1

RESULT 8
US-08-462-040-22/c
Sequence 22, Application US/08462040
Patent No. 6177554
GENERAL INFORMATION:
APPLICANT: Moo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gotchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,040
FILING DATE: June 5, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-462-040-22

Query Match 61.8%; Score 13.6; DB 3; Length 27;
Best Local Similarity 80.0%; Pred. NO. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCCCTCCGACCCACA 20
DB 20 CTCCTCCCTCCCTCCCA 1

RESULT 9
US-09-396-196G-58648/c
Sequence 58648, Application US/09396196G
Patent No. 6821724
GENERAL INFORMATION:
APPLICANT: Michael Miltmann
APPLICANT: David Lockhart
APPLICANT: Affymetrix, Inc.
TITLE OF INVENTION: Methods of Genetic Analysis
FILE REFERENCE: 3101.1
CURRENT APPLICATION NUMBER: US/09/396,196G
CURRENT FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: 60/100,678
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 127806
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58648
LENGTH: 25
TYPE: DNA
ORGANISM: mus musculus
US-09-396-196G-58648

Query Match 60.9%; Score 13.4; DB 3; Length 25;
Best Local Similarity 93.3%; Pred. NO. 1.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCCTCCAGCCACATG 22
DB 24 GCCTCCAGCCACAG 10

RESULT 10
US-09-396-196G-58649/c
Sequence 58649, Application US/09396196G
Patent No. 6821724
GENERAL INFORMATION:
APPLICANT: Michael Miltmann
APPLICANT: David Mack
APPLICANT: David Lockhart
APPLICANT: Affymetrix, Inc.
TITLE OF INVENTION: Methods of Genetic Analysis
FILE REFERENCE: 3101.1
CURRENT APPLICATION NUMBER: US/09/396,196G
CURRENT FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: 60/100,678
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 127806
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58649
LENGTH: 25
TYPE: DNA
ORGANISM: mus musculus
US-09-396-196G-58649

Query Match 60.9%; Score 13.4; DB 3; Length 25;
Best Local Similarity 93.3%; Pred. NO. 1.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCCTCCAGCCACATG 22
DB 18 GCCTCCAGCCACAG 4

RESULT 11
US-09-396-196G-58650/c
Sequence 58650, Application US/09396196G
Patent No. 6821724
GENERAL INFORMATION:
APPLICANT: Michael Miltmann
APPLICANT: David Mack
APPLICANT: David Lockhart
APPLICANT: Affymetrix, Inc.
TITLE OF INVENTION: Methods of Genetic Analysis
FILE REFERENCE: 3101.1
CURRENT APPLICATION NUMBER: US/09/396,196G
CURRENT FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: 60/100,678
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 127806
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58650
LENGTH: 25
TYPE: DNA
ORGANISM: mus musculus
US-09-396-196G-58650

Query Match 60.9%; Score 13.4; DB 3; Length 25;
Best Local Similarity 93.3%; Pred. NO. 1.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCCTCCAGCCACATG 22
DB 15 GCCTCCAGCCACAG 1

RESULT 12
US-08-766-982-10
Sequence 10, Application US/08766982
Patent No. 5948892
GENERAL INFORMATION:
APPLICANT: Wahl, Robert C.
TITLE OF INVENTION: Analogs of Macrophage Stimulating
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,982
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-441
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

US-08-766-982-10

Query Match 59.1%; Score 13; DB 2; Length 23;

Best Local Similarity 76.2%; Pred. No. 2.1e+04;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCCCTCGCCTCCAGCCACATG 22

Db 2 TGCTTCTCCTCCTCGCCATATG 22

RESULT 13

US-09-296-219-10

Sequence 10, Application US/09296219

Patent No. 6248560

GENERAL INFORMATION:

APPLICANT: Wahl, Robert C.

TITLE OF INVENTION: Analogs of Macrophage Stimulating

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: 1840 Dehaven Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/296,219

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Winter, Robert B.

REFERENCE/DOCKET NUMBER: A-441

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 23 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-296-219-10

Query Match 59.1%; Score 13; DB 3; Length 23;

Best Local Similarity 76.2%; Pred. No. 2.1e+04;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCCCTCGCCTCCAGCCACATG 22

Db 2 TGCTTCTCCTCCTCGCCATATG 22

RESULT 14

US-08-529-190B-16

Sequence 16, Application US/08529190B

Patent No. 5833991

GENERAL INFORMATION:

APPLICANT: Masucci, Maria G.

TITLE OF INVENTION: GLYCINE-CONTAINING SEQUENCES

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.

STREET: One Financial Center

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02111

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/529,190B

FILING DATE: 15-SEP-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: SE9501324-9

FILING DATE: 10-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US08/522,595

FILING DATE: 01-SEP-1995

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Ph.D., Kathleen A

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 3555/53015

TELEPHONE: 617-345-9100

TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 24 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

US-08-529-190B-16

Query Match 59.1%; Score 13; DB 2; Length 24;

Best Local Similarity 76.2%; Pred. No. 2.1e+04;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACAT 21

Db 4 CACCCGACCTCCAGCTCCAT 24

RESULT 15

US-09-119-507B-10

Sequence 10, Application US/09119507B

Patent No. 6548642

GENERAL INFORMATION:

APPLICANT: Kieliszewski, Marcia J.

TITLE OF INVENTION: No. 6548642el Synthetic Genes for Plant Gums

FILE REFERENCE: OHU-03417

CURRENT APPLICATION NUMBER: US/09/119,507B

CURRENT FILING DATE: 1998-07-20

NUMBER OF SEQ ID NOS: 118

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO: 10

LENGTH: 30

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-119-507B-10

Query Match 59.1%; Score 13; DB 3; Length 30;

Best Local Similarity 76.2%; Pred. No. 2.1e+04;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACAT 21

Db 5 CACCTCAGCTCCAGCCCAT 25

Search completed: March 25, 2006, 13:03:00

Job time : 129.5 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

- Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
- Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).



This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelleraion Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 13:59:41 ; Search time 660.5 Seconds
(without alignments)
275.437 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22

Sequence: 1 CTCCTCGCCTCCAGCCACATG 22

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 10535742

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	8	US-10-798-652-7
2	20.4	92.7	22	8	US-10-798-652-6
3	15.8	71.8	25	10	US-11-036-317-589884
4	15.6	70.9	25	10	US-11-036-317-179644
5	15.6	70.9	25	10	US-11-036-317-250940
6	15.6	70.9	25	10	US-11-036-317-253919
7	15.6	70.9	25	10	US-11-036-317-268146
8	15.6	70.9	25	10	US-11-036-317-321184
9	15.6	70.9	25	10	US-11-036-317-391184
10	15.6	70.9	25	10	US-11-060-756-199930
11	15.4	70.0	20	9	US-10-994-138-20
12	15.2	69.1	25	10	US-11-036-317-278200
13	15.2	69.1	25	10	US-11-036-317-286177
14	14.6	66.4	25	10	US-11-036-317-280075
15	14.6	66.4	25	10	US-11-036-317-360952
16	14.6	66.4	25	10	US-11-036-317-469308
17	14.6	66.4	25	10	US-11-036-317-674241
18	14.6	66.4	25	10	US-11-060-756-20600
19	14.6	66.4	25	10	US-11-060-756-20618
20	14.4	65.5	25	10	US-11-036-317-889096
21	14.4	65.5	25	10	US-11-060-756-236805
22	14.4	65.5	25	10	US-11-060-756-236806
23	14.2	64.5	25	8	US-10-719-900-693282

C 24	14.2	64.5	25	9	US-10-956-157-212800	Sequence 212800,
C 25	14.2	64.5	25	9	US-10-956-157-252303	Sequence 252303,
C 26	14.2	64.5	25	9	US-10-956-157-257894	Sequence 257894,
C 27	14.2	64.5	25	9	US-10-956-157-257895	Sequence 257895,
C 28	14.2	64.5	25	10	US-11-036-317-70925	Sequence 70925, A
C 29	14.2	64.5	25	10	US-11-036-317-343201	Sequence 343201, A
C 30	14.2	64.5	25	10	US-11-036-317-589885	Sequence 589885,
C 31	14.2	64.5	25	10	US-11-036-317-846119	Sequence 846119,
C 32	14	63.6	25	7	US-10-681-773-22338	Sequence 22338, A
C 33	14	63.6	25	7	US-10-681-773-223893	Sequence 72893, A
C 34	14	63.6	25	9	US-10-485-048-86	Sequence 86, Appl
C 35	14	63.6	25	9	US-10-956-157-174568	Sequence 174568,
C 36	14	63.6	25	9	US-10-956-157-250478	Sequence 250478,
C 37	14	63.6	25	10	US-11-036-317-516736	Sequence 516736,
C 38	14	63.6	25	10	US-11-036-317-923714	Sequence 923714,
C 39	13.8	62.7	24	3	US-09-940-185-956	Sequence 956, App
C 40	13.8	62.7	25	8	US-10-719-900-16414	Sequence 16414, A
C 41	13.8	62.7	25	9	US-10-956-157-163759	Sequence 163759,
C 42	13.8	62.7	25	10	US-11-036-317-516615	Sequence 516615,
C 43	13.8	62.7	25	10	US-11-036-317-579814	Sequence 579814,
C 44	13.8	62.7	25	10	US-11-036-317-579816	Sequence 579816,
C 45	13.6	61.8	25	5	US-10-215-112-12019	Sequence 12019, A

ALIGNMENTS

```
RESULT 1
US-10-798-652-7
; Sequence 7, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICANT: Yongjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; FILE REFERENCE: 3382-PO3136US01
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/445,698
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: probe
US-10-798-652-7

Query Match      100.0%; Score 22; DB 8; Length 22;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACATG 22
Db 1 CTCCTCGCCTCCAGCCACATG 22

RESULT 2
US-10-798-652-6
; Sequence 6, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICANT: Yongjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; FILE REFERENCE: 3382-PO3136US01
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/445,698
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 7
```

```
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 6
/ LENGTH: 22
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: probe
US-10-798-652-6

Query Match          92.7%; Score 20.4; DB 8; Length 22;
Best Local Similarity 95.5%; Pred. No. 57;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACATG 22
Db 1 CTCCTCCACCTCCAGCCACATG 22

RESULT 3
US-11-036-317-589884/c
/ Sequence 589884, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 589884
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-589884

Query Match          71.8%; Score 15.8; DB 10; Length 25;
Best Local Similarity 89.5%; Pred. No. 4.8e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCCTCGCCTCCAGCCACAT 21
Db 25 CCCTCGCCTCCAGCCCTCT 7

RESULT 4
US-11-036-317-179644/c
/ Sequence 179644, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 179644
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-179644

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.9e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 CTCCTCGCCTCCAGCCACATG 22
Db 23 CCCACTCCCTCCAGCCACAG 2

RESULT 5
US-11-036-317-250940
/ Sequence 250940, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 250940
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-250940

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.9e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACATG 22
Db 2 CTCCTCCACATGAGCCACATG 23

RESULT 6
US-11-036-317-253919/c
/ Sequence 253919, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
/ FILE REFERENCE: 3654.1
/ CURRENT APPLICATION NUMBER: US/11/036,317
/ CURRENT FILING DATE: 2005-01-13
/ PRIOR APPLICATION NUMBER: US 60/536,639
/ PRIOR FILING DATE: 2004-01-13
/ NUMBER OF SEQ ID NOS: 991174
/ SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
/ SEQ ID NO 253919
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-11-036-317-253919

Query Match          70.9%; Score 15.6; DB 10; Length 25;
Best Local Similarity 81.8%; Pred. No. 5.9e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCCTCGCCTCCAGCCACATG 22
Db 22 CCCACTCCCTCCAGCCACAG 1

RESULT 7
US-11-036-317-268615
/ Sequence 268615, Application US/11036317
/ Publication No. US20050214823A1
/ GENERAL INFORMATION:
/ APPLICANT: Williams, Alan
/ APPLICANT: Blume, John
/ TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
```

FILE REFERENCE: 3654.1
CURRENT APPLICATION NUMBER: US/11/036,317
CURRENT FILING DATE: 2005-01-13
PRIOR APPLICATION NUMBER: US 60/536,639
NUMBER OF SEQ ID NOS: 991174
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO: 268615
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-11-036-317-268615

Query Match
Best Local Similarity 81.8%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACATG 22
Db 1 CTCCTCCCATGCAGCCACCTG 22

RESULT 8
US-11-036-317-362146
Sequence 362146, Application US/11036317
Publication No. US20050214823A1
GENERAL INFORMATION:
APPLICANT: Williams, Alan
APPLICANT: Blume, John
TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
FILE REFERENCE: 3654.1
CURRENT APPLICATION NUMBER: US/11/036,317
CURRENT FILING DATE: 2005-01-13
PRIOR APPLICATION NUMBER: US 60/536,639
PRIOR FILING DATE: 2004-01-13
NUMBER OF SEQ ID NOS: 991174
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO: 362146
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-11-036-317-362146

Query Match
Best Local Similarity 81.8%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACATG 22
Db 3 CTCCTCCCATGCAGCCACCTG 24

RESULT 9
US-11-036-317-391184
Sequence 391184, Application US/11036317
Publication No. US20050214823A1
GENERAL INFORMATION:
APPLICANT: Williams, Alan
APPLICANT: Blume, John
TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
FILE REFERENCE: 3654.1
CURRENT APPLICATION NUMBER: US/11/036,317
CURRENT FILING DATE: 2005-01-13
PRIOR APPLICATION NUMBER: US 60/536,639
PRIOR FILING DATE: 2004-01-13
NUMBER OF SEQ ID NOS: 991174
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO: 391184
LENGTH: 25
TYPE: DNA
ORGANISM: Mus musculus
US-11-036-317-391184

Query Match
Best Local Similarity 81.8%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACATG 22
Db 4 CTCCTCCCATGCAGCCACCTG 25

RESULT 10
US-11-060-756-199930/c
Sequence 199930, Application US/11060756
Publication No. US20050221354A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Mount, William Martin
TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
FILE REFERENCE: AM101083 (031896-042000)
CURRENT APPLICATION NUMBER: US/11/060,756
CURRENT FILING DATE: 2005-02-18
NUMBER OF SEQ ID NOS: 303284
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 199930
LENGTH: 25
TYPE: DNA
ORGANISM: probe
US-11-060-756-199930

Query Match
Best Local Similarity 81.8%; Score 15.6; DB 10; Length 25;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCCTCGCCTCCAGCCACATG 22
Db 25 CTCCTCGCCCCCGACACAGG 4

RESULT 11
US-10-994-138-20
Sequence 20, Application US/10994138
Publication No. US2005018666A1
GENERAL INFORMATION:
APPLICANT: Dow Global Technologies, Inc.
TITLE OF INVENTION: Improved Protein Expression Systems
FILE REFERENCE: 00588.105020
CURRENT APPLICATION NUMBER: US/10/994,138
CURRENT FILING DATE: 2004-11-19
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 20
LENGTH: 20
TYPE: DNA
ORGANISM: artificial
FEATURE:
OTHER INFORMATION: pyranalr
US-10-994-138-20

Query Match
Best Local Similarity 94.1%; Score 15.4; DB 9; Length 20;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CTCGCTTCAGCCACAT 21
Db 1 CTCGCTTCAGCCACAT 17

RESULT 12
US-11-036-317-278200/c
Sequence 278200, Application US/11036317
Publication No. US20050214823A1
GENERAL INFORMATION:
APPLICANT: Williams, Alan

```

; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 278200
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-278200

```

```

Query Match          69.1%; Score 15.2; DB 10; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.6e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      1 CTCCTCGGCTCCAGCCACA 20
Db      21 CCCACTCCCTCCAGCCACA 2

```

```

RESULT 13
US-11-036-317-386177/c
; Sequence 386177, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 386177
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-386177

```

```

Query Match          69.1%; Score 15.2; DB 10; Length 25;
Best Local Similarity 85.0%; Pred. No. 8.6e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      1 CTCCTCGGCTCCAGCCACA 20
Db      20 CCCACTCCCTCCAGCCACA 1

```

```

RESULT 14
US-11-036-317-280075
; Sequence 280075, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 280075
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus

```

```

US-11-036-317-280075

```

```

Query Match          66.4%; Score 14.6; DB 10; Length 25;
Best Local Similarity 81.0%; Pred. No. 1.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      1 CTCCTCGGCTCCAGCCACATG 21
Db      5 CTCCTCCCATGAGCCACCT 25

```

```

RESULT 15
US-11-036-317-360952
; Sequence 360952, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 360952
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-360952

```

```

Query Match          66.4%; Score 14.6; DB 10; Length 25;
Best Local Similarity 81.0%; Pred. No. 1.5e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      2 TCCTCGGCTCCAGCCACATG 22
Db      1 TCCTCCCATGAGCCACCTG 21

```

```

Search completed: March 25, 2006, 15:50:32
Job time : 661.5 secs

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 14:11:31 ; Search time 315.5 Seconds
(without alignments)
277.964 Million cell updates/sec

Title: US-10-798-652-7

Perfect score: 22
Sequence: 1 cccctgcgcctccagcacatg 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9258654 seqs, 1993127192 residues

Total number of hits satisfying chosen parameters: 13041728

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA New:*
1: /SIDS5/pcodaca/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/pcodaca/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/pcodaca/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/pcodaca/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/pcodaca/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/pcodaca/1/pubpna/US09_NEW_PUB.seq1:*
7: /SIDS5/pcodaca/1/pubpna/US10_NEW_PUB.seq1:*
8: /SIDS5/pcodaca/1/pubpna/US10_NEW_PUB.seq2:*
9: /SIDS5/pcodaca/1/pubpna/US10_NEW_PUB.seq3:*
10: /SIDS5/pcodaca/1/pubpna/US10_NEW_PUB.seq4:*
11: /SIDS5/pcodaca/1/pubpna/US11_NEW_PUB.seq2:*
12: /SIDS5/pcodaca/1/pubpna/US11_NEW_PUB.seq3:*
13: /SIDS5/pcodaca/1/pubpna/US11_NEW_PUB.seq4:*
14: /SIDS5/pcodaca/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/pcodaca/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.8	71.8	20	8	US-10-310-914A-613434 Sequence 613434
2	15.8	71.8	21	8	US-10-310-914A-148615 Sequence 148615
3	15.8	71.8	23	8	US-10-310-914A-743419 Sequence 743419
4	15.4	70.0	22	8	US-10-310-914A-743458 Sequence 743458
5	15.2	69.1	20	8	US-10-310-914A-290068 Sequence 290068
6	15.2	69.1	21	8	US-10-310-914A-177473 Sequence 177473
7	15.2	69.1	21	8	US-10-310-914A-613446 Sequence 613446
8	15.2	69.1	24	8	US-10-310-914A-233742 Sequence 233742
9	15.2	69.1	24	8	US-10-310-914A-549830 Sequence 549830
10	15.2	69.1	25	8	US-10-310-914A-549844 Sequence 549844
11	15.2	69.1	22	8	US-10-310-914A-165922 Sequence 165922
12	14.8	67.3	18	8	US-10-310-914A-213337 Sequence 213337
13	14.8	67.3	19	8	US-10-310-914A-120157 Sequence 120157
14	14.8	67.3	19	8	US-10-310-914A-213338 Sequence 213338
15	14.8	67.3	19	8	US-10-310-914A-668834 Sequence 668834
16	14.8	67.3	19	8	US-10-310-914A-668992 Sequence 668992
17	14.8	67.3	20	8	US-10-310-914A-213339 Sequence 213339
18	14.8	67.3	20	8	US-10-310-914A-317759 Sequence 317759

C 19	14.8	67.3	20	8	US-10-310-914A-377554	Sequence 377554
C 20	14.8	67.3	20	8	US-10-310-914A-701328	Sequence 701328
C 21	14.8	67.3	20	8	US-10-310-914A-1138593	Sequence 1138593
C 22	14.8	67.3	20	14	US-11-093-746A-14	Sequence 14, Appl
C 23	14.8	67.3	21	8	US-10-310-914A-1108644	Sequence 1108644
C 24	14.8	67.3	21	8	US-10-310-914A-1138594	Sequence 1138594
C 25	14.8	67.3	25	14	US-11-121-849-439265	Sequence 439265
C 26	14.6	66.4	24	8	US-10-310-914A-237510	Sequence 237510
C 27	14.6	66.4	25	8	US-10-310-914A-91862	Sequence 91862, A
C 28	14.6	66.4	25	8	US-10-310-914A-91863	Sequence 91863, A
C 29	14.6	66.4	29	8	US-10-857-780-2372	Sequence 2372, Ap
C 30	14.4	65.5	20	8	US-10-310-914A-532861	Sequence 532861
C 31	14.4	65.5	21	8	US-10-310-914A-102269	Sequence 102269
C 32	14.4	65.5	21	8	US-10-310-914A-102291	Sequence 102291
C 33	14.4	65.5	21	8	US-10-310-914A-532862	Sequence 532862
C 34	14.4	65.5	21	8	US-10-310-914A-782776	Sequence 782776
C 35	14.4	65.5	22	8	US-10-310-914A-782816	Sequence 782816
C 36	14.4	65.5	25	9	US-10-933-982-36071	Sequence 36071, A
C 37	14.2	64.5	19	8	US-10-310-914A-554165	Sequence 554165
C 38	14.2	64.5	19	8	US-10-310-914A-448406	Sequence 448406
C 39	14.2	64.5	20	8	US-10-310-914A-519040	Sequence 519040
C 40	14.2	64.5	20	8	US-10-310-914A-1002041	Sequence 1002041
C 41	14.2	64.5	20	8	US-10-310-914A-1223911	Sequence 1223911
C 42	14.2	64.5	20	8	US-10-310-914A-1352507	Sequence 1352507
C 43	14.2	64.5	21	8	US-10-310-914A-724597	Sequence 724597
C 44	14.2	64.5	21	8	US-10-310-914A-1352505	Sequence 1352505
C 45	14.2	64.5	21	8	US-10-310-914A-1352505	Sequence 1352505

ALIGNMENTS

RESULT 1
US-10-310-914A-613434
; Sequence 613434, Application US/10310914A
; Publication No. US2006000322A1
; GENERAL INFORMATION:
; APPLICANT: Benitich, Isaac
; TITLE OF INVENTION: Biointernationally detectable group of novel regulatory genes an
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 613434
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Human
; US-10-310-914A-613434

Query Match 71.8%; Score 15.8; DB 8; Length 20;
Best Local Similarity 73.7%; Pred. No. 1.9e+03;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCCTGCGCTCCAGCCACAT 21
DB 2 CCCTGACCCUCCAGCCACAU 20

RESULT 2
US-10-310-914A-148615
; Sequence 148615, Application US/10310914A
; Publication No. US2006000322A1
; GENERAL INFORMATION:
; APPLICANT: Benitich, Isaac
; TITLE OF INVENTION: Biointernationally detectable group of novel regulatory genes and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06

```
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 148615
/ LENGTH: 21
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-148615
```

```
Query Match          71.8%; Score 15.8; DB 8; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.9e+03;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CTCCTCGCCTCCAGCCAC 19
      |||:|||||:|||||
Db      1 CUCCGCGCCGCUCCAGCCUC 19
```

```
RESULT 3
US-10-310-914A-743419/c
/ Sequence 743419, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kyuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 743419
/ LENGTH: 23
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-743419
```

```
Query Match          71.8%; Score 15.8; DB 8; Length 23;
Best Local Similarity 89.5%; Pred. No. 1.9e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CTCCTCGCCTCCAGCCAC 19
      |||:|||||:|||||
Db      22 CTCCTCGCCTCCAGCCAC 4
```

```
RESULT 4
US-10-310-914A-743458/c
/ Sequence 743458, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kyuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 743458
/ LENGTH: 22
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-743458
```

```
Query Match          70.0%; Score 15.4; DB 8; Length 22;
Best Local Similarity 94.1%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 CTCCTCGCCTCCAGCC 17
      |||:|||||:|||||
Db      18 CTCCTCGCCTCCAGCC 2
```

```
RESULT 5
US-10-310-914A-290068
/ Sequence 290068, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kyuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 290068
/ LENGTH: 20
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-290068
```

```
Query Match          69.1%; Score 15.2; DB 8; Length 20;
Best Local Similarity 70.0%; Pred. No. 3.4e+03;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 CTCCTCGCCTCCAGCCACA 20
      |||:|||||:|||||
Db      1 CUCCTCCTCCAGCCACA 20
```

```
RESULT 6
US-10-310-914A-177473
/ Sequence 177473, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kyuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ NUMBER OF SEQ ID NOS: 1388402
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 177473
/ LENGTH: 21
/ TYPE: RNA
/ ORGANISM: Human
US-10-310-914A-177473
```

```
Query Match          69.1%; Score 15.2; DB 8; Length 21;
Best Local Similarity 80.0%; Pred. No. 3.4e+03;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 CTCCTCGCCTCCAGCCACA 20
      |||:|||||:|||||
Db      1 CUCCTCCTCCAGCCACA 20
```

```
RESULT 7
US-10-310-914A-631446
/ Sequence 631446, Application US/10310914A
/ Publication No. US20060003322A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Bentwich, Isaac
/ APPLICANT: Shlier, Kyuzat
/ TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
/ FILE REFERENCE: 06087.0200.CPUS01
/ CURRENT APPLICATION NUMBER: US/10/310,914A
/ NUMBER OF SEQ ID NOS: 1388402
```

```

; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3..

```

SEQ ID NO 213337
LENGTH: 18
TYPE: RNA
ORGANISM: Human
US-10-310-914A-213337

Query Match 67.3%; Score 14.8; DB 8; Length 18;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTGCGCTCCAGCCA 18
DB 18 CTCCTGCGCTCCAGTCA 1

RESULT 13
US-10-310-914A-120157/C
Sequence 120157, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shiller, Kiyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 120157
LENGTH: 19
TYPE: RNA
ORGANISM: Human
US-10-310-914A-120157

Query Match 67.3%; Score 14.8; DB 8; Length 19;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCTGCGCTCCAGCCAC 19
DB 19 TCCTGCGCTCCAGCCAC 2

RESULT 14
US-10-310-914A-213338/C
Sequence 213338, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shiller, Kiyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 213338
LENGTH: 19
TYPE: RNA
ORGANISM: Human
US-10-310-914A-213338

Query Match 67.3%; Score 14.8; DB 8; Length 19;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCCTGCGCTCCAGCCA 18
DB 18 CTCCTGCGCTCCAGTCA 1

RESULT 15
US-10-310-914A-668834/C
Sequence 668834, Application US/10310914A
Publication No. US20060003322A1
GENERAL INFORMATION:
APPLICANT: Bentwich, Isaac
APPLICANT: Shiller, Kiyazat
TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
FILE REFERENCE: 06087.0200.CPUS01
CURRENT APPLICATION NUMBER: US/10/310,914A
CURRENT FILING DATE: 2002-12-06
NUMBER OF SEQ ID NOS: 1388402
SOFTWARE: PatentIn version 3.3
SEQ ID NO 668834
LENGTH: 19
TYPE: RNA
ORGANISM: Human
US-10-310-914A-668834

Query Match 67.3%; Score 14.8; DB 8; Length 19;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCTGCGCTCCAGCCACA 20
DB 18 CCTGCGCTCCAGCCACA 1

Search completed: March 25, 2006, 14:54:05
Job time : 316.5 secs

QY 121 GGGGCTTGGCCATGTTGGGACACAGGGGCTGCTGATACACTTACTGATGACATTCAGTCC 180
| | | | |
DB 51638 GGGGCTTGGCCATGTTGGGACACAGGGGCTGCTGATACACTTACTGATGACATTCAGTCC 51579
| | | | |
QY 181 CTGGTCCAGGCTCTGGAAGTCTGGAGAGGACATGTTTCCATTAAAGAAAGTGTG 240
| | | | |
DB 51578 CTGGTCCAGGCTCTGGAAGTCTGGAGAGGACATGTTTCCATTAAAGAAAGTGTG 51519
| | | | |
QY 241 GCGGCGCATGCCCCCAACGTTGCACTACTGCTTGTGCAAGGTTGGGGCTTCCAGTCC 300
| | | | |
DB 51518 GCGGCGCATGCCCCCAACGTTGCACTACTGCTTGTGCAAGGTTGGGGCTTCCAGTCC 51459
| | | | |
QY 301 ACAGGCTCCCATTCACAGTACCGAGCCAGGTGCTGCAAGAGTCCCTCGCACTGAAA 360
| | | | |
DB 51458 ACAGGCTCCCATTCACAGTACCGAGCCAGGTGCTGCAAGAGTCCCTCGCACTGAAA 51399
| | | | |
QY 361 CCAAGGGAAGGCTTGGGAAAACCATCTGAAGGGACATGGCTTGAATTTAGTGAAGAGGTGG 420
| | | | |
DB 51398 CCAAGGGAAGGCTTGGGAAAACCATCTGAAGGGACATGGCTTGAATTTAGTGAAGAGGTGG 51339
| | | | |
QY 421 GCGTGGGCTGGGCAAGGCAAGGCTGATGATGAGAGGCAAGGCAAGGCAAGGCTGCTCC 480
| | | | |
DB 51338 GCGTGGGCTGGGCAAGGCAAGGCTGATGATGAGAGGCAAGGCAAGGCAAGGCTGCTCC 51279
| | | | |
QY 481 A-GCACTGCGCGCGCTCTGCTGATGAGTCTCTGCGCACCTGAGAAAGCTGTAGA 539
| | | | |
DB 51278 AGGCACTGCGCGCGCTCTGCTGATGAGTCTCTGCGCACCTGAGAAAGCTGTAGA 51219
| | | | |
QY 540 GAGGAGTGGGCTCTTTGGGACTTC 564
| | | | |
DB 51218 GAGGAGTGGGCTCTTTGGGACTTC 51194
| | | | |

RESULT 2
US-09-925-065A-734471

/ Sequence 734471, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925, 065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243, 096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252, 147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250, 092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261, 766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289, 846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 734471
/ LENGTH: 541
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-734471

Query Match 7.1%; Score 39.8; DB 6; Length 541;
Best Local Similarity 64.8%; Pred. No. 0.64;
Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 4 GCGCTGCTCAGAAAAACAGAGAGCAGACACTCAGCGTACTACCCCATGTGCT 63
| | | | |
DB 442 GCGCTGCTCAGAAAAACAGAGAGCAGACACTCAGCGTACTACCCCATGTGAA 501
| | | | |
QY 64 GAGAGCGAGAGAGCTCTCTGAGGCAAGGCCA 94
| | | | |
DB 502 TGATGAGAGGTGTTCCTCCCAAGGCTGGGCCA 532
| | | | |

RESULT 3
US-11-121-086-20/c
/ Sequence 20, Application US/11121086
/ Publication No. US20050266459A1
/ GENERAL INFORMATION:
/ APPLICANT: POULSEN, TIM S.
/ APPLICANT: NIELSEN, KIRSTEN V.
/ TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
/ FILE REFERENCE: 09138.6000-00000
/ CURRENT APPLICATION NUMBER: US/11/121, 086
/ CURRENT FILING DATE: 2005-05-04
/ PRIOR APPLICATION NUMBER: 60/567, 570
/ PRIOR FILING DATE: 2004-05-04
/ NUMBER OF SEQ ID NOS: 107
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 20
/ LENGTH: 162289
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-121-086-20

Query Match 6.7%; Score 38; DB 14; Length 162289;
Best Local Similarity 46.6%; Pred. No. 5.2;
Matches 122; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 3 AGCCCTGCTCAGAAAAACAGAGAGCAGACACTCAGCGTACTACCCCATGTGCT 62
| | | | |
DB 145401 AGCTCAGCAGACACTCAGCGAGGAGCCACCTCTGCTGTGAGTGAATCAGAGAG 145342
| | | | |
QY 63 TGAAGCGAGGAGCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 122
| | | | |
DB 145341 CTAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145282
| | | | |
QY 123 GCTCTGCAATGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 182
| | | | |
DB 145281 GTCATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145222
| | | | |
QY 183 GGTGCGAGCTCTGGAAGTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 242
| | | | |
DB 145221 GAGACACTTGGAGACTGACAGACACTGTGGGGGCTCTGGAGCGCAAGCAACTAAGTCTC 145162
| | | | |
QY 243 CGGCCATGCCCCCAACGCTTGC 264
| | | | |
DB 145161 ACAACAACACCCCGAGGCTCTGC 145140
| | | | |

RESULT 4
US-10-750-185-32987
/ Sequence 32987, Application US/10750185
/ Publication No. US20050260603A1
/ GENERAL INFORMATION:
/ APPLICANT: MMI GENOMICS, INC.
/ APPLICANT: DENISE, Sue K.
/ APPLICANT: KERR, Richard
/ APPLICANT: ROSENFELD, David
/ APPLICANT: HOLM, Tom
/ APPLICANT: BATES, Stephen

/ APPLICANT: PANTIN, Dennis
/ TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

/ FILE REFERENCE: MM1100-2
/ CURRENT APPLICATION NUMBER: US/10/750, 185
/ CURRENT FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437, 482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 32987
/ LENGTH: 712
/ TYPE: DNA

/ ORGANISM: Bovine 19866880739882
US-10-750-185-32987

Query Match 6.6%; Score 37; DB 8; Length 712;
Best Local Similarity 47.6%; Pred. No. 3.8;
Matches 109; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

```
QY 307 TCCATCCACGTAACCAAGCCAGGCTGTCAGAGAGTCCCTCCGCTCATGAACCAAG 366
DB 283 TGCCGACACATTCAGTCTTATGGAAAGATCAATAATTAAGAGATGAATAATGCTG 342
QY 367 GAGCCTGGGAAACCAATCGAAGGCGATGGCTTTGATTAAGAGAGGGGGGGCTGG 426
DB 343 AAAAAGTGTAAAGCAATATTAAGCAATTATGCTTATGTAAGAGAGGGGTCTCT 402
QY 427 GCTGGGCAAGGCCACAGGTCTGAGTCAAGGCCAGAGGCAAGAGCTGCTCCAGCACT 486
DB 403 TGCTTACTCCCTTACTGGAATAATAGCGAAGGCCAGGATGTCAGAGATCCTCAAGCT 462
QY 487 GCGGCGCCCTCTGTCGATGACGCTCTCTGGCCACTGAAGACGCTG 535
DB 463 TGCTTCATCTCTGTCTTCATTTCTCAATTAAGAGCCTCGAAGAGACTG 511
```

RESULT 5
US-10-750-623-32987
Sequence 32987, Application US/10750623

Publication No. US20050287531A1
GENERAL INFORMATION:

APPLICANT: MMT GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Denise
TITLE OR INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-1
CURRENT FILING DATE: 2003-12-31
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 32987
LENGTH: 712
TYPE: DNA
ORGANISM: Bovine 19866880739882
US-10-750-623-32987

Query Match 6.6%; Score 37; DB 8; Length 712;
Best Local Similarity 47.6%; Pred. No. 3.8;
Matches 109; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

```
QY 307 TCCATCCACGTAACCAAGCCAGGCTGTCAGAGAGTCCCTCCGCTCATGAACCAAG 366
DB 283 TGCCGACACATTCAGTCTTATGGAAAGATGAATAATTAAGAGATGAATAATGCTG 342
QY 367 GAGCCTGGGAAACCAATCGAAGGCGATGGCTTTGATTAAGAGAGGGGGGGCTGG 426
DB 343 AAAAAGTGTAAAGCAATATTAAGCAATTATGCTTATGTAAGAGAGGGGTCTCT 402
QY 427 GCTGGGCAAGGCCACAGGTCTGAGTCAAGGCCAGAGGCAAGAGCTGCTCCAGCACT 486
DB 403 TGCTTACTCCCTTACTGGAATAATAGCGAAGGCCAGGATGTCAGAGATCCTCAAGCT 462
QY 487 GCGGCGCCCTCTGTCGATGACGCTCTCTGGCCACTGAAGACGCTG 535
DB 463 TGCTTCATCTCTGTCTTCATTTCTCAATTAAGAGCCTCGAAGAGACTG 511
```

RESULT 6
US-09-925-065A-57042/c
Sequence 57042, Application US/09925065A
Publication No. US20040181048A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OR INVENTION: Identification and Mapping of Single
TITLE OR INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 57042
LENGTH: 1172
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-57042

Query Match 6.5%; Score 36.4; DB 6; Length 1172;
Best Local Similarity 59.8%; Pred. No. 6;
Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY 25 GAGCAGACACTACAGTGTACTACCCCATGTGCTGAGAGCGAGAGCCTCTGA 84
DB 599 GGGGCGAGCCCTGCGATGTGACTTTTGGCAGATGAGAGAGACACAGAGCAGCA 540
QY 85 GGCAGGGCCAGGCGAGCCTGTAGTGGGTGACGCGAGGGTC 126
DB 539 GGAAGGAGAGCGCAGCCAGCAGCAGAGGTGACAGGTGACATC 498
```

RESULT 7

US-10-301-480-158280/c
Sequence 158280, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OR INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301, 480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215, 598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311, 695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 158280
LENGTH: 1172
TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-158280

Query Match 6.5%; Score 36.4; DB 9; Length 1172;
Best Local Similarity 59.8%; Pred. No. 6;
Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY 25 GAGCAGACACTACAGTGTACTACCCCATGTGCTGAGAGCGAGAGCCTCTGA 84
DB 599 GGGGCGAGCCCTGCGATGTGACTTTTGGCAGATGAGAGAGACACAGAGCAGCA 540
QY 85 GGCAGGGCCAGGCGAGCCTGTAGTGGGTGACGCGAGGGTC 126
DB 539 GGAAGGAGAGCGCAGCCAGCAGCAGAGGTGACAGGTGACATC 498
```

```

RESULT 8
US-10-301-480-771689/c
; Sequence 771689, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 771689
; LENGTH: 1172
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-771689

```

Query Match	6.5%	Score 36.4;	DB 10;	Length 1172;
Best Local Similarity	59.8%	Pred. No. 6;		
Matches 61;	Conservative	0;	Mismatches 41;	Indels 0;
				Gaps 0.

QY 25 GAGACGACACCTCAACGTACTCACCCCATGTGGCTGAGAGCGAGAGCTCTCGA 84
Db 599 GGGGGAGCCCTTGCGCATGTGACTTTGGCAGATGGAGAGAAAGACACAGAGAGCAGAGA 540

Oy 85 GGCAGGGCCAGGGCAGCGCTCAGTGGGTGACGGCAGGGTC 126
 ||||| | | ||||| ||| ||||| | ||
Db 539 GGAAAGGAAGAGCGCACGCCACCGCAGGTGACAGGTGACATC 498

```

RESULT 9
US-10-995-561-63329/C
; Sequence 63329, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001359
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63329
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-63329

```

Query Match	6.3%	Score 35.8;	DB 8;	Length 201;
Best Local Similarity	56.6%	Pred. No. 6.5;		
Matches	64;	Conservative	1;	Mismatches 48;
				Indels 0;
				Gaps 0

QY 45 GACTACACCCCATGTCGCTGGAGGCCAGGAGACCTCCTGAGGCAGAGGCCAGGCGCT 104
 138 GACTACCCGTCCTGGCTCTGCGTGTGCGCGCCGCGCCAGGCTGTGTCACAGACCTGG 79
 Db

QY 105 CAGGTGGTGACGCGCAGGGCTTGGCCATGTGGGCACAGGGGCTGCATACAG 157
||||| | ||||| | ||||| | ||||| | ||||| |
Db 78 TGGGTGCGGGTGGGGCTGGCTTCCTGCCATCAGGCGCACGGTGGCTTCATGCAG 26

RESULT 10
US-10-995-561-69610
; Sequence 69610, Application US/10995561

```

; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69610
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-69610

```

Query Match	6.3%	Score 35.8;	DB 8;	Length 201;
Best Local Similarity	56.6%	Pred. No. 6.5;		
Matches 64;	Conservative 1;	Mismatches 48;	Indels 0;	Gaps 0;

QY		43	GACTTACCCCCCATGTGGCTGAGGCGCAGGAGCCCTCTGAGGCAGGCCAAGGCGCAAGCCGT	104
Db		64	GACTTACCCGTCGTGGCTCTGCTGCTGATGTTGGCCGCGCCGACAGGCTGTGTTCACGACGTGG	123

Oy 105 CAGTGGGTGAACGCGAGGGTTCTTGCCATGGTGGGCACAAGGGGCTTCATACAG 157
 ||||| | ||||| | ||||| |
Db 124 TGGGTGCCGGTGGGGCTGGCTCCTGCATCAGSGCACGGTGGCTTCAATGCAG 176

```

RESULT 11
US-10-995-561-13412
/ Sequence 13412, Application US/10995561
/ Publication No. US20050272054A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
/ TITLE OF INVENTION: DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01559
/ CURRENT APPLICATION NUMBER: US/10/995,561
/ CURRENT FILING DATE: 2004-11-24
/ NUMBER OF SEQ ID NOS: 85702
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13412
/ LENGTH: 25257
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-995-561-13412

```

Query Match	6.3%	Score 35.8;	DB 0;	Length 25257;
Best Local Similarity	56.6%	Pred. No. 15;		
Matches 64;	Conservative 1;	Mismatches 48;	Indels 0;	Gaps 0;

QY 45 GACTCACCCCATGTGCTGGAGCGCAGGAGACCTTCTTGAGCAGGCGCAGGGCAGCCTT 104
Db 24045 GACTCACCGTGTCTTGTGCTCTGTGCATGTGCCCGGCAGGCTGTGTTCAGCAGCTGG 24106	

QY 105 CAGGTGGGTGACGCGAGGGGCTTGCATGTGTGGGACAGGGGCTGCATTACAG 157
|||||
24105 TGGGTGCGGGGTGGGGCTGGCTCTTCGCATCAGGCGGCACGGTGGCTTCATGCAG 24157

```

RESURT 12
US-10-995-561-13364/C
; Sequence 13364, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559

```


RESULT 14
US-10-301-480-93596/c
; Sequence 93596, Application US/10301480
; Publication No. US20060057564A1

Query Match	6.2%	Score 35;	DB 10;	Length 1741;
Best Local Similarity	51.6%	Pred. No. 15;		
Matches	80;	Conservative	0;	Mismatches 75; Indels 0; Gaps 0

QY	72	GGAGCCTCTTGAGGCAGGGCCAGGGCAGCCGTCAGTGGGTATACGGACGGGCTTTGCC	131
DB	190	GGGATGATGCTGACACAGGGAGCCTTGAGGGGGCTGGGAGAGGGCAGGGAAAGGGCTCTCCCC	131
QY	132	ATGCTGGGCACAGGGGCTGCATACAGCTTACTCAGTACCAATCGAGTCCCTGGTGCACG	191
DB	130	AAAGCCTTCAGGGAGCATGGTCTCTGTGACATCTTGAACGACAGACATCGCTTCCACA	71
QY	192	CTCTGAAGTCTGCAAGTAGCAATGTTTCCATT	226

Db 70 GTATGACAGAGTCATTTCGGTGAAGTGCCTCAGT 36

Search completed: March 25, 2006, 18:31:06
Job time : 441.5 secs

QY	121	GGGGCTCTTGGCACTGGTGGGACAGAGGGCTCATACAGCTTACTCAGTGCACATCGAGTCC	180
Db	51638	GGGGTCTTGGCACTGGTGGGACAGAGGGCTCATACAGCTTACTCAGTGCACATCGAGTCC	51579
QY	181	CTGGTGCACGCTCTTGGAGCTTGGAGTGGAGCAATGTTTCCCATTAAGGAAAGTGTGTG	240
Db	51578	CTGGTGCACGCTCTTGGAGAGTGGAGTGGAGCAATGTTTCCCATTAAGGAAAGTGTGTG	51519
QY	241	GGCGGCGCATGCCCCCGCCAAAGTTGCACACTACGCTCTTGGCAGAGTTGGGGCTTCCAGTCC	300
Db	51518	GCCGCGCATGCCCCCGCCAAAGTTGCACACTACGCTCTTGGCAGAGTTGGGGCTTCCAGTCC	51458
QY	301	ACAGAGTCCCATCCATCCAGTACCAAGCCAGTGTGGCTGCAAGAGTCCCTCGCAGTCATGAAG	360
Db	51458	ACAGAGTCCCATCCATCCAGTACCAAGCCAGTGTGGCTGCAAGAGTCCCTCGCAGTCATGAAG	51399
QY	361	CCAAAGGAGGCTTGGGAAACCAACTCTGAAGGCGATGGCTTTGATATTGATGAGGGGTGG	420
Db	51398	CCAAAGGAGGCTTGGGAAACCAACTCTGAAGGCGCGATGGCTTTGATATTGATGAGAGGGGTGG	51339
QY	421	GGCTGGGCTGGGGCAAGGCGCACAGAGTGTGAGTCAAGACCAAGAGCGAGAAAGTGTGTCCC	480
Db	51338	GGCTGGGCTGGGGCAAGGCGCACAGAGTGTGAGTCAAGACCAAGAGCGAGAAAGTGTGTCCC	51279
QY	481	A-GCAGTGCCTCGCGCGCTCTTGCAGTGCAGTCTTCTGGCCACTGTGAACACAGCTGTAGA	539
Db	51278	AGGCACTGCCTCGCGCGCTCTTGCAGTGCAGTCTTCTGGCCACTGTGAACACAGCTGTAGA	51219
QY	540	GAGGCACTGCAGTCTTTCGGAATTC	564
Db	51218	GAGGCACTGCAGTCTTTCGGAATTC	51194

```

RESULT 2
US-09-925-065A-734471
Sequence 734471, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-06-08
PRIORITY APPLICATION NUMBER: US 60/243,096
PRIORITY FILING DATE: 2000-10-24
PRIORITY APPLICATION NUMBER: US 60/252,147
PRIORITY FILING DATE: 2000-11-20
PRIORITY APPLICATION NUMBER: US 60/250,092
PRIORITY FILING DATE: 2000-11-30
PRIORITY APPLICATION NUMBER: US 60/261,766
PRIORITY FILING DATE: 2001-01-16
PRIORITY APPLICATION NUMBER: US 60/289,846
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 734471
LENGTH: 541
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-734471

```

Qy	4	GCCCTGCGCTCGAAGAAAGAGAGAGACACATCACGCTGATCTACATCCCATGTGGCT	63
Db	442	GCCCTCCCTCGGAAACACAGAGAGGCTGGGGACCCCTCGTGGGCTACCCCTCATTTGGAA	501
Qy	64	GGAGGTAGGGAGAGCCTCTCTGAGCAGAGGCCA	94
Db	502	TGATGAGAGGTGTTCCCAAGGCTCGGGGCA	532

```

RESULT 3
US-11-121-086-20/c
; Sequence 20, Application US/11121086
; Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S
APPLICANT: NIELSEN, KRISTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121.086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatencIn version 3.3
SEQ ID NO 20
LENGTH: 162289
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-20

```

	Query Match	7.0%;	Score 39.6;	DB 14;	Length 162289;
	Best Local Similarity	46.9%;	Pred. No.2,1;		
	Matches	123;	Conservative	0;	Mismatches 139;
				Indels	0;
				Gaps	0;
Oy	3	AGCCCTGCTCAGAAAAACAGAAAGCAGACACACTCAGCGTGACTCACCCCATTTGGC	62		
Db	145401	AGCTCAGAGAGAACCTTGGGGAGGGCCACTCTGTGGCTCTGGATGGGATCACAGAAAG	145342		
Oy	63	TGAGAGTGAAGGAGCCTCTGAGGCGAGGGCAGGGCAGCCCTCAGATGGGTGACGACAG	122		
Db	145341	CCTAGGTAGGGAGGAGCTGGAAGCCAGGGCCCTCAGAAGAGTGGGGTTGGGCAATTGGGGAG	145282		
Oy	123	GCTCTTGCCATGATGGGGCACAAGGGGCTGCATACGCTTACTAATGATGACAAATGAGTCCCT	182		
Db	145281	GTCATTGGCAGAGTGGGGCAGGTCAAGGCTGTGAGCACCTTCCCAAGGCCAATGACCAAAAT	145222		
Oy	183	GGTCCAGACCTCTGGAAAGTGTGAAAGTAGCATGTTTCCATTAAAGAAAGTGTGTGC	242		
Db	145221	GAGACAACTTTGGGACTGACAGAGCATGTGGGGGGCTCTTGGAGCCAGAAAGCTAAATGTTC	145162		
Oy	243	CGGCCATGCCCCCAAGCTTGC	264		
Db	145161	ACACAAACACCCCAAGCTTGC	145140		

```

RESULT 4
US-10-750-185-32987
Sequence 32987, Application US/10750185
Publication No. US200502603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM11100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 32987
LENGTH: 712
TYPE: DNA
ORGANISM: Bovine
US-10-750-185-32987
19866880739882

```

Query Match	6.6%	Score 37	DB 8	Length 712
Best Local Similarity	47.6%	Pred. No. 4.1		
Matches 109, Conservative	0	Mismatches 120	Indels 0	Gaps 0

Qy	30	TCCCATCTCAGTACAGAGCCAGGTCCTCACAAGAGTCCCTGCAGTCACTAAACCAAGG	366
Db	283	TGCTGACACATTCACATCCTATTGGGAGATGAAATTAATTAGAGACATGAAAAATGCTG	342
Qy	367	GAGGCTTGGGAAAACAATCTGAAGGCGATGGCTTTGATTGATGAGAGGCTGGGGCTGG	428
Db	343	AAAAATGGTAAAGACAGATATAAAGCATTTATCTTACTGTATGAAAGAGTGGGTCTCT	402
Qy	427	GCTGGGCAAGCCACAAGCTTGAGCTGAGTCAGAGCCAGAGGACAGGAGCTGCTCCCAACACT	488
Db	403	TGCTTACTCCCTCACTGGAAAAATGAGCGAAGGCGAGTGATCCCAAGGACATCTCAGACT	462
Qy	487	GCCGCGCGCTCTGCATGCAGTGCCTCTCTGGCGACCTGAGAAACAGCTTG	535
Db	463	TGCTCATCCTCTGTCTTCAATTCCTCAATTAAGGCTCGGAGAGACCTG	511

RESULT 5
US-10-750-623-32987

```

Sequence 32987, Application US/10750623
Publication No. US20050287531A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM11100-1
CURRENT APPLICATION NUMBER: US/10/750,623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64932
SOFTWARE: Patentin version 3.1
SEQ ID NO 32987
LENGTH: 712
TYPE: DNA
ORGANISM: Bovine 19866880739882
US-10-750-623-32987

```

Query Match	6.6%	Score 37;	DB 8;	Length 712;
Best local Similarity	47.6%	Pred. NO. 4.1;		
Matches 109; Conservative	0;	Mismatches 120;	Indels 0;	Gaps 0

Qy	307	TCCATCTCAAGTACCAAGCCAGAGTGGCTGCAAAAGTCCCTGCAGTCAATAAAACAAGG	366
Db	283	TGCTTACACATTCACTGCTCTTAATTGGAGAGATGAAATTAATTAGAGCATGAAATAATCTG	342
Qy	367	GAGGCTTGGGAAACCACTATGTAGAGGGCATGCTTGATTTAGTGAAGAGGTGGGGCTGG	428
Db	343	AAAAGTGTAAAGCAGATATATAAGCATTAATGCTTACTGTGATGAAAGAGTGGTCTCT	402
Qy	427	GCTGGCAAGGCCACAGGTTGAGTCAAGCCAGAGGCAAGGCAAGGCTGGTCCCAAGACT	488
Db	403	TGCTTACTCCCTACCTGGGAAATAGGCGAAGGCCAGTGAATCCAAAGCATCTCAAGACT	462
Qy	487	GCCGCGCGGCTCTGCGATGCAAGTCTCTCTGGCGAATCTAGAAACAGCTCG	535
Db	463	TGCTCATCTCTGTCTTCCATTCTCCATTAAGAGCTCTGGAAAGAGACTG	511

RESULT 6
US-10-995-561-63329/c
; Sequence 63329, Application US/10995561
; Publication No. US20050272054A1

GENERAL INFORMATION: Michele et al.
APPLICANT: CARGILL,
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CU001559

Query Match	6.3%;	Score 35.8;	DB 8;	length 201;
Best Local Similarity	56.6%;	Pred. No. 6.9;		
Matches 64;	Conservative 1;	Mismatches 48;	Indels 0;	Gaps 0;

Qy	4	GACTCA	CCCCC	ATG	TGG	CTG	GAG	GGT	GAA	GAG	AGC	CTT	CGA	AGG	CA	GAG	CA	GCG	CGT	104
Db	138	GACTCA	CGGT	GT	TGT	GCT	CTG	GGT	TGG	CGC	CGC	KAG	CGT	GT	TTT	CA	GAG	CA	GCT	79
Qy	105	CAG	CGT	CGT	GAA	GGG	CA	GAG	GGT	CTT	CGC	AT	TGG	CA	GAG	GGC	GT	CA	TAT	157
Db	78	TGG	GT	TGG	GT	GGG	CGT	GGC	CTT	CGC	AT	CA	GAG	GGC	CA	CGG	GT	TAT	GAG	26

RESULT 7
US-10-995-561-69610

```

: Sequence 69610, Application US/10995561
: Publication No. US20050272054A1
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
: TITLE OF INVENTION: DETECTION AND USES THEREOF
: FILE REFERENCE: CL001559
: CURRENT APPLICATION NUMBER: US/10/995,561
: CURRENT FILING DATE: 2004-11-24
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 69610
: LENGTH: 201
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-995-561.69610

```

Query Match	6.34;	Score 35.8;	DB 8;	Length 201;
Best Local Similarity	56.64;	Pred. No. 6.9;		
Matches	64;	Conservative	1;	Mismatches 48; Indels 0; Gaps 0;

QY 45 GACTACCCCCCATGTTGGCTGTGAGAGGTGAAGAGAGCCCTCCCTACAGACGGGCGAGGGACGGCGT 10
Db 64 GACTCACCGTCGTTGTGTCTCTGCGTGGTGGCCGCCCKAGGCTGTGTGTTCCAGACAGCTGG 12
QY 105 CAGGTGGGTGAACGGCAGGGGCTTTTCCATGTTGGGCACAGGGGCTGCTATACAG 157
Db 124 TGGAGTCCGGGTGGGGCTGGCTTCCTCCATCAGAGGGCAGCGAGTGCTTATATGAG 176

RESULT 8
 US-10-995-561-13412
 ; Sequence 13412, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; TITLE OF INVENTION: DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001359
 ; CURRENT APPLICATION NUMBER: US/10/995.561


```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,655
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 707005
; LENGTH: 1741
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-707005

```

Query Match	6.2%	Score 35;	DB 10;	Length 1741;
Best Local Similarity	51.6%	Pred. No. 16;		
Matches	80;	Conservative	0;	Mismatches 75; Indels 0; Gaps 0;
Qy	72	GGAGGCTCTGAGGCAGGGCCAGGACCCTCAGTGGGTGACGGCAGGGCTTTC	131	
Db	190	GGGATGATGCTGCACACAGGGAGCGTGGAGGGCTGGAGAGGCAGGAAGGTCCTCCC	131	
Qy	132	ATGTGGGACACAGGGGCTGCATACAGCTTACTCAGTACAATCGAGTCCCTGGTCCAGC	191	
Db	130	AGAGCTTCAGGGGGGCAATGCTCCCTGCATCTTTGATCTGCAGACATCTGGCTTCAG	71	
Qy	192	CTCTGGAAGTCTGGAAGTGAACAATGTTCCATT	226	
Db	70	GTATGACAGAGTGCATTTCCGTGAAGTGCAGCCAGT	36	

```

RESULT 13
US-10-750-185-42100/c
; Sequence 42100, Application US/10750185
; Publication No. US2005026063A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OR INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750.185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42100
; LENGTH: 835
; TYPE: DNA
; ORGANISM: Bovine 19866880864778
; US-10-750-185-42100

```

	Query Match	Similarity	Score	DB	Length
Best Local	84	50.6%	6.24	8	835
Matches	84	Conservative	0	Mismatches	82
				Indels	0
				Gaps	0

	Query	Match	Similarity	Score	DB	Length
164	CAGTGACAAATGATATCCCTGTGTGCGACAGCCTCTTGGAATCTGGAATGATGCAATGTTTCCC	223				
538	CAGCAACAGACGCGTCCCTGATGTCAGTAATGGCCAGACAGCAGTGGATGAAACGTG	479				
224	ATTAAAGAAATGTGTGTGGCCGCGCAATGCCCCCAACGTTGCACTCACTGCTTTTGAG	283				
478	GGCGAGGAGATGTGCTCTCCCTGGAGATGACATCAAGTTCAGGCCCAATCAATCACTGC	419				

```
Oy      284 GGTGGGGCTTCCAGTACAGGGGTCCCATCCAGTACCAGCCCAAG 329
          |||||
Db      418 GGCTGTGGCACCCAGCCCAACTCTAACTCTCTCCATCAG 373
```

```

RESULT 14
US-10-750-623-42100/C
Sequence 42100, Application US/10750623
Publication No. US20050287531a1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Denise
TITLE OR INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM11100-1
CURRENT APPLICATION NUMBER: US/10/750, 623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42100
LENGTH: 835
TYPE: DNA
ORGANISM: Bovine 19866880864778
US-10-750-623-42100

```

	Query Match	Best Local Similarity	Score 34.8;	DB 8;	Length 835;
	Matches	84;	Conservative	0;	Mismatches 82;
					Indels 0;
					Gaps 0
Qy	164	CAGTGAACATTCAGATCCCTGTGTGCAGCCTTCGAACTGTGAAGTGAAGCATGTTTCCC	223		
Db	538	CAGCAACAGACGCGTCCCTGATGTACAGTATATGGCAGACGACGTGGAGATAGGAACGTG	479		
Qy	224	ATTAAAGAAAGTGTGTGTGGCGGCATGCCCCCAACGTTGCACACTACTGCCCTTTGCAG	283		
Db	478	GGCAGGGGAAGGTGTGCTGTGCTCGGATGAGACTCAACAGCTCAGCCCAATCAATCACTGCAA	419		
Qy	284	GGTTGGGGCTTCCAGTCAACAGGGGTCCATCCACATCAAGCCCAAG	329		
Db	418	GGCTGTGTGGACCCAGCCCAACTCTTCAACATCTCTCTCCATCAAG	373		

```

RESULT 15
US-09-925-065A-57042/C
: Sequence 57042, Application US/09925065A
: Publication No. US20040181048A1
: GENERAL INFORMATION:
: APPLICANT: Mang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single
: FILE REFERENCE: 108827.135
: CURRENT APPLICATION NUMBER: US/09/925, 065A
: CURRENT FILING DATE: 2001-08-08
: PRIOR APPLICATION NUMBER: US 60/243, 096
: PRIOR FILING DATE: 2000-10-24
: PRIOR APPLICATION NUMBER: US 60/252, 147
: PRIOR FILING DATE: 2000-11-20
: PRIOR APPLICATION NUMBER: US 60/250, 092
: PRIOR FILING DATE: 2000-11-30
: PRIOR APPLICATION NUMBER: US 60/261, 766
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/289, 846
: PRIOR FILING DATE: 2001-05-09
: NUMBER OF SEQ ID NOS: 957086
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 57042
: LENGTH: 1172

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-57042

Query Match 6.2%; Score 34.8; DB 6; Length 1172;
Best Local Similarity 58.8%; Pred. No. 17;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Oy 25 GGACCGACGACACTCAGCGTGACTCAGCCCATGTGCTGAGGTGAGGAGGAGCTTCTGA 84
|||
Db 599 GGGGGCAGCCCTGGCATGTGTGCTTGGCAGATGAGAGAGACACAGAGCAGCGAGA 540
Oy 85 GGCAGGGCCAGGGCAGCCGCTCAGGTGGGTGACGCGCAGGGGTC 126
|||
Db 539 GGAAGGAGAGCGCGCAGCCACAGCGAGGTGACAGGTGACATC 498
|||

Search completed: March 25, 2006, 18:31:07
Job time : 441.5 secs


```
Db      241 GCGGCGCAATGCCCGCCCAAGTTCACATCTGCTTTGCAAGGTTGGGCTTCCAGTC 300
Qy      301 ACAGGCTCCATCCACGTAACAGCCAGGTTGAGAGAGGCTCCGCAATGATAA 360
Db      301 ACAGGCTCCATCCATGACAGCCAGGTTGAGAGAGGCTCCGCAATGATAA 360
Qy      361 CCAGGGAAGCTTGGGAAACCATCTGAAGGGCATGGCTTTAGTTAGTAGAGGTTG 420
Db      361 CCAGGGAAGCTTGGGAAACCATCTGAAGGGCATGGCTTTAGTTAGTAGAGGTTG 420
Qy      421 GGGTGGGCTGGGCAAGGCCACAGGTTGATGATGAGGCCAGAGGCAAGCTGTCCTCC 480
Db      421 GGGTGGGCTGGGCAAGGCCACAGGTTGATGATGAGGCCAGAGGCAAGCTGTCCTCC 480
Qy      481 AGCACTGCCCGCGGCTCTGCGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db      481 AGCACTGCCCGCGGCTCTGCGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy      541 AGGCACTGGCGCTCTTTCGAGCTTC 564
Db      541 AGGCACTGGCGCTCTTTCGAGCTTC 564
```

```
RESULT 2
US-10-027-632-142127
; Sequence 142127, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142127
```

```
Query Match      99.6%; Score 562; DB 5; Length 700;
Best Local Similarity 99.6%; Pred. No. 6,4e-164;
Matches 562; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 GCAGCCCTGCTCAGAAAACAGAAAGCAGACACATCAGGTGACTACCCCATGTG 60
Db      107 GCAGCCCTGCTCAGAAAACAGAAAGCAGACACATCAGGTGACTACCCCATGTG 166
Qy      61 GCTGAGGTGAGGAGCCTCTTGAAGGCGGCAAGGCGGCTCAAGTGGTGAAGGCA 120
Db      167 GCTGAGGTGAGGAGCCTCTTGAAGGCGGCAAGGCGGCTCAAGTGGTGAAGGCA 226
Qy      121 GGGGTCTTGCAATGATGAGGCAAGGCGGCTCAATGAGCTTCAATGAGCTGTC 180
Db      227 GGGGTCTTGCAATGATGAGGCAAGGCGGCTCAATGAGCTTCAATGAGCTGTC 286
```

```
Qy      181 CTGTGCCAGCCTCTGGAAGTCTGGAAGTGAACAAATGTTTCCATTAAAGAAAGTGTG 240
Db      287 CTGTGTGCCAGCCTCTGGAAGTCTGGAAGTGAACAAATGTTTCCATTAAAGAAAGTGTG 346
Qy      241 GCGGCGCAATGCCCGCCCAAGTTCACATCTGCTTTGCAAGGTTGGGCTTCCAGTC 300
Db      347 GCGGCGCAATGCCCGCCCAAGTTCACATCTGCTTTGCAAGGTTGGGCTTCCAGTC 406
Qy      301 ACAGGCTCCATCCACGTAACAGCCAGGTTGAGAGAGGCTCCGCAATGATAA 360
Db      407 ACAGGCTCCATCCACGTAACAGCCAGGTTGAGAGAGGCTCCGCAATGATAA 466
Qy      361 CCAGGGAAGCTTGGGAAACCATCTGAAGGGCATGGCTTTAGTTAGTAGAGGTTG 420
Db      467 CCAGGGAAGCTTGGGAAACCATCTGAAGGGCATGGCTTTAGTTAGTAGAGGTTG 526
Qy      421 GGGTGGGCTGGGCAAGGCCACAGGTTGATGATGATGATGATGATGATGATGATGATG 480
Db      527 GGGTGGGCTGGGCAAGGCCACAGGTTGATGATGATGATGATGATGATGATGATGATG 586
Qy      481 AGCACTGCCCGCGGCTCTGCGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db      587 AGCACTGCCCGCGGCTCTGCGATGATGATGATGATGATGATGATGATGATGATGATG 646
Qy      541 AGGCACTGGCGCTCTTTCGAGCTTC 564
Db      647 AGGCACTGGCGCTCTTTCGAGCTTC 670
```

```
RESULT 3
US-10-027-632-142127
; Sequence 142127, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142127
```

```
Query Match      99.6%; Score 562; DB 6; Length 700;
Best Local Similarity 99.6%; Pred. No. 6,4e-164;
Matches 562; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 GCAGCCCTGCTCAGAAAACAGAAAGCAGACACATCAGGTGACTACCCCATGTG 60
Db      107 GCAGCCCTGCTCAGAAAACAGAAAGCAGACACATCAGGTGACTACCCCATGTG 166
Qy      61 GCTGAGGTGAGGAGCCTCTTGAAGGCGGCAAGGCGGCTCAAGTGGTGAAGGCA 120
Db      167 GCTGAGGTGAGGAGCCTCTTGAAGGCGGCAAGGCGGCTCAAGTGGTGAAGGCA 226
```

QY 121 GGGGCTTTCATGCTGGGCAAGGGGCTGATACAGCTTACTAGTGAATTCAGTCC 180
 DB 227 GGGGCTTTCATGCTGGGCAAGGGGCTGATACAGCTTACTAGTGAATTCAGTCC 286
 QY 181 CTGGTCCAGGCTCTGGAAGTCTGGAAGTGAAGCAATGTTCCATTAAAGAAAGTGTG 240
 DB 287 CTGGTCCAGGCTCTGGAAGTCTGGAAGTGAAGCAATGTTCCATTAAAGAAAGTGTG 346
 QY 241 GCGGCGCATGCCCCCAAGTTCACATCAGCTTTCAGGGGTGGGGCTTCAGTCC 300
 DB 347 GCGGCGCATGCCCCCAAGTTCACATCAGCTTTCAGGGGTGGGGCTTCAGTCC 406
 QY 301 ACAGGCTCCATCCATGATACAGGCTGAGTGGCTGAGAGGTTCCCTGCACTATGAAA 360
 DB 407 ACAGGCTCCATCCATGATACAGGCTGAGTGGCTGAGAGGTTCCCTGCACTATGAAA 466
 QY 361 CCAAGGAGGCTTGGGAAACCAATCTGAAAGGCAATGCTTTGATTAGTGAAGAGTGG 420
 DB 467 CCAAGGAGGCTTGGGAAACCAATCTGAAAGGCAATGCTTTGATTAGTGAAGAGTGG 526
 QY 421 GGGTGGGCTGGGCAAGGCAAGGCTGAGTCAAGGCAAGGCAAGGCAAGGCTGGTCCC 480
 DB 527 GGGTGGGCTGGGCAAGGCAAGGCTGAGTCAAGGCAAGGCAAGGCAAGGCTGGTCCC 586
 QY 481 AGCACTGCGCGCGCTCTGAGTCAAGTCTCTGAGCACTGAGAAAGGCTGTAGAG 540
 DB 587 AGCACTGCGCGCGCTCTGAGTCAAGTCTCTGAGCACTGAGAAAGGCTGTAGAG 646
 QY 541 AGGCACTGCGCTCTTTCGAGCTTC 564
 DB 647 AGGCACTGCGCTCTTTCGAGCTTC 670

RESULT 4
 US-10-027-632-142126
 ; Sequence 142126, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 142126
 ; LENGTH: 855
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-10-027-632-142126

Query Match 98.0%; Score 552.6; DB 5; Length 855;
 Best Local Similarity 99.6%; Pred. No. 5.4e-161;
 Matches 563; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 GCAGCCCTGCTCAGAAAAAGAGAGCAGACACTCAGGTGACTCACTCCCATGTG 60

DB 108 GCAAGCTTCCTCAGAAAAAGAGAGCAGACACTCAGGTGACTCACTCCCATGTG 167
 QY 61 GCTGAGGTGAGAGAGCTCTCTGAGGCAAGGCAAGGCAAGGCAAGGCTGAGTGA 120
 DB 168 GCTGAGGTGAGAGAGCTCTCTGAGGCAAGGCAAGGCAAGGCAAGGCTGAGTGA 227
 QY 121 GGGGCTTTCATGCTGGGCAAGGGGCTGATACAGCTTACTAGTGAATTCAGTCC 180
 DB 227 GGGGCTTTCATGCTGGGCAAGGGGCTGATACAGCTTACTAGTGAATTCAGTCC 287
 QY 181 CTGGTCCAGGCTCTGGAAGTCTGGAAGTGAAGCAATGTTCCATTAAAGAAAGTGTG 240
 DB 287 CTGGTCCAGGCTCTGGAAGTCTGGAAGTGAAGCAATGTTCCATTAAAGAAAGTGTG 347
 QY 241 GCGGCGCATGCCCCCAAGTTCACATCAGCTTTCAGGGGTGGGGCTTCAGTCC 300
 DB 347 GCGGCGCATGCCCCCAAGTTCACATCAGCTTTCAGGGGTGGGGCTTCAGTCC 407
 QY 301 ACAGGCTCCATCCATGATACAGGCTGAGTGGCTGAGAGGTTCCCTGCACTATGAAA 360
 DB 407 ACAGGCTCCATCCATGATACAGGCTGAGTGGCTGAGAGGTTCCCTGCACTATGAAA 467
 QY 361 CCAAGGAGGCTTGGGAAACCAATCTGAAAGGCAATGCTTTGATTAGTGAAGAGTGG 420
 DB 467 CCAAGGAGGCTTGGGAAACCAATCTGAAAGGCAATGCTTTGATTAGTGAAGAGTGG 527
 QY 421 GGGTGGGCTGGGCAAGGCAAGGCTGAGTCAAGGCAAGGCAAGGCAAGGCTGGTCCC 480
 DB 527 GGGTGGGCTGGGCAAGGCAAGGCTGAGTCAAGGCAAGGCAAGGCAAGGCTGGTCCC 587
 QY 481 A-GCACTGCGCGCGCTCTGAGTCAAGTCTCTGAGCACTGAGAAAGGCTGTAGAG 539
 DB 588 AGCACTGCGCGCGCTCTGAGTCAAGTCTCTGAGCACTGAGAAAGGCTGTAGAG 647
 QY 540 AGGCACTGCGCTCTTTCGAGCTTC 564
 DB 648 AGGCACTGCGCTCTTTCGAGCTTC 672

RESULT 5
 US-10-027-632-142126
 ; Sequence 142126, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 142126
 ; LENGTH: 855
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-10-027-632-142126

Query Match 98.0%; Score 552.6; DB 6; Length 855;
 1 GCAGCCCTGCTCAGAAAAAGAGAGCAGACACTCAGGTGACTCACTCCCATGTG 60

Best Local Similarity 99.6%; Pred. No. 5,4e-161; Matches 563; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```
OY 1 GCAGCCCTGCTCAGAAAACAGAAAGGACGACACATCAGCGTGACTCACCCCATG 60
Db 108 GCAGCCCTGCTCAGAAAACAGAAAGGACGACACATCAGCGTGACTCACCCCATG 167
OY 61 GCTGAGGTGAGGAGACCTCTGAGGCAAGGCGCCAGGCGCTGCAGGTGAGCGCA 120
Db 168 GCTGAGGTGAGGAGACCTCTGAGGCAAGGCGCCAGGCGCTGCAGGTGAGCGCA 227
OY 121 GGGGCTTTGGCANTGTGGGCAAGGGGCTGACATGCTTACTCACTGACATTCAGTCC 180
Db 228 GGGGCTTTGGCANTGTGGGCAAGGGGCTGACATGCTTACTCACTGACATTCAGTCC 287
OY 181 CTGGTCCAGCCCTGGAAGTCTGAAAGTGAAGTATTTCCCATTAAGAAAGTGTG 240
Db 288 CTGGTCCAGCCCTGGAAGTCTGAAAGTGAAGTATTTCCCATTAAGAAAGTGTG 347
OY 241 GCTGGGCAATGCCCCCAACGTTGCACTCACTGCTTTGCAAGGTTGGGCTTCCAGTC 300
Db 348 GCTGGGCAATGCCCCCAACGTTGCACTCACTGCTTTGCAAGGTTGGGCTTCCAGTC 407
OY 301 AAGAGGTCCCATTCACATGACAGCCAGGTGCTGCAAGAGTCCCTGCGCATGAGAA 360
Db 408 AAGAGGTCCCATTCACATGACAGCCAGGTGCTGCAAGAGTCCCTGCGCATGAGAA 467
OY 361 CCAAGGAGGCTTGGGAAACCACTGAAAGGCGATGGCTTTGATTGAGAGGGGTG 420
Db 468 CCAAGGAGGCTTGGGAAACCACTGAAAGGCGATGGCTTTGATTGAGAGGGGTG 527
OY 421 GGTGGGCTGGGCAAGGCCACAGGTCTGATCAGAGCCAGAGGCAAGGCTGTGCCC 480
Db 528 GGTGGGCTGGGCAAGGCCACAGGTCTGATCAGAGCCAGAGGCAAGGCTGTGCCC 587
OY 481 A-GCACTGCCCCCGCTCTGCGANTGACATCTCTCTGGGCACTGTGAAACACCTGTAA 539
Db 588 AAGCACTGCCCCCGCTCTGCGANTGACATCTCTCTGGGCACTGTGAAACACCTGTAA 647
OY 540 GAGGCAGTGGCGCTTTCCGACTTC 564
Db 648 GAGGCAGTGGCGCTTTCCGACTTC 672
```

RESULT 6
US-10-322-281-642
; Sequence 642, Application US/10322281
; Publication No. US20040126762A1

GENERAL INFORMATION:
APPLICANT: David W. Morris
APPLICANT: Marc S. Malandro
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 529452001000
CURRENT APPLICATION NUMBER: US/10/322,281
PRIOR FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 866
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 642
; LENGTH: 29340
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-642

Query Match 97.8%; Score 551.4; DB 7; Length 29340;
Best Local Similarity 99.6%; Pred. No. 2.1e-160;

Matches 563; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
OY 1 GCAGCCCTGCTCAGAAAACAGAAAGGACGACACATCAGCGTGACTCACCCCATG 60
Db 3566 GCAGCCCTGCTCAGAAAACAGAAAGGACGACACATCAGCGTGACTCACCCCATG 3625
OY 61 GCTGAGGTGAGGAGACCTCTGAGGCAAGGCGCCAGGCGCTGCAGGTGAGCGCA 120
```

Db 3626 GCTGAGGCGAGGAGACCTCTGAGGCAAGGCGCCAGGCGCTGCAGGTGAGCGCA 3685

OY 121 GGGGCTTTGGCANTGTGGGCAAGGGGCTGACATGCTTACTCACTGACATTCAGTCC 180

Db 3686 GGGGCTTTGGCANTGTGGGCAAGGGGCTGACATGCTTACTCACTGACATTCAGTCC 3745

OY 181 CTGGTCCAGCCCTGGAAGTCTGAAAGTGAAGTATTTCCCATTAAGAAAGTGTG 240

Db 3746 CTGGTCCAGCCCTGGAAGTCTGAAAGTGAAGTATTTCCCATTAAGAAAGTGTG 3805

OY 241 GCGGCGCATGCCCCCAACGTTGCACTCACTGCTTTGCAAGGTTGGGCTTCCAGTC 300

Db 3806 GCGGCGCATGCCCCCAACGTTGCACTCACTGCTTTGCAAGGTTGGGCTTCCAGTC 3865

OY 301 AAGAGGTCCCATTCACATGACAGCCAGGTGCTGCAAGAGTCCCTGCGCATGAGAA 360

Db 3866 AAGAGGTCCCATTCACATGACAGCCAGGTGCTGCAAGAGTCCCTGCGCATGAGAA 3925

OY 361 CCAAGGAGGCTTGGGAAACCACTGAAAGGCGATGGCTTTGATTGAGAGGGGTG 420

Db 3926 CCAAGGAGGCTTGGGAAACCACTGAAAGGCGATGGCTTTGATTGAGAGGGGTG 3985

OY 421 GGTGGGCTGGGCAAGGCCACAGGTCTGATCAGAGCCAGAGGCAAGGCTGTGCCC 480

Db 3986 GGTGGGCTGGGCAAGGCCACAGGTCTGATCAGAGCCAGAGGCAAGGCTGTGCCC 4045

OY 481 A-GCACTGCCCCCGCTCTGCGANTGACATCTCTCTGGGCACTGTGAAACACCTGTAA 539

Db 4046 AAGCACTGCCCCCGCTCTGCGANTGACATCTCTCTGGGCACTGTGAAACACCTGTAA 4105

OY 540 GAGGCAGTGGCGCTTTCCGACTTC 564

Db 4106 GAGGCAGTGGCGCTTTCCGACTTC 4130

RESULT 7
US-09-925-065A-734471
; Sequence 734471, Application US/09925065A
; Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 734471
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-734471

Query Match 7.1%; Score 39.8; DB 4; Length 541;
Best Local Similarity 64.8%; Pred. No. 0.062; Indels 0; Gaps 0;

Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
OY 4 GCCCTGCTCAGAAAACAGAAAGGACGACACATCAGCGTGACTCACCCCATG 63
Db 442 GCCCTGCTCAGAAAACAGAAAGGACGACACCTCTCGTGGCTCACCCTCATTTGAA 501
OY 64 GAGGTGAGGAGACCTCTGAGGCAAGGCGCA 94
```

Db 502 TGATGAGAGGTGTTCCCGCAGGCTGGGCCA 532

RESULT 8

US-10-108-260A-578/c
Sequence 578, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1el full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 578
LENGTH: 2777
TYPE: DNA
ORGANISM: Homo sapiens
US-10-108-260A-578

Query Match 6.8%; Score 38.2; DB 6; Length 2777;
Best Local Similarity 48.8%; Pred. No. 0.25;

Matches 103; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 79 TCCTGAGCAGGCGCCAGGCGCAGGTGAGCGGCGGCTTGCATGTTGG 138
Db 345 TCCGCCCCAGAGGTGTCGCCACGCTGGGGGGTCACTGCGGCGCAAGGCTGGGGGG 286
Qy 139 GCACAGGCGCTGCATACAGCTTACTCAGTGCATCGATCCCTGCTGCGCAGCTCTGGA 198
Db 285 CCACCGGAGCCCCACCTGCGGTGTCACCCGAGCAGTCGTTGGGCCCCGCGCACTCAC 226
Qy 199 AGTCTGAGAGTACGAAAGTTTCCCATTTAGGAAAGTGTGTGGCCGCGCATGCCCCCAA 258
Db 225 AGTGAAGCAGAGGCACTTGTTCTCTCTGATGACCCAGGCGCGCCAGGAGCCGATGAC 166
Qy 259 CGTGCACAGCTCAGTCCCTTGCAGAGGTTGG 289
Db 165 CATGACAAAGGCGCGGTGATGATGAGCAGG 135

RESULT 9

US-10-123-155-10
Sequence 10, Application US/10123155
Publication No. US20030068794A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Aubert L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10

LENGTH: 594
TYPE: PRT
ORGANISM: Homo sapien
US-10-123-155-10

Query Match 6.7%; Score 37.8; DB 5; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.26;

Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

Qy 23 AAGACGACGACACTCAGGTGACTACCCCCATGTGGCTGAGAGTAGAGGAGCTTCT 82
Db 56 VALAAESPVMWMPQGDNNMIDRPVAAHDPYPLPITISPEQSDERKCYNER 115
Qy 83 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 142
Db 116 YRGIVQNDPAGISEQCLYQIYIBLYGLORPSEDEKKLAEEKASIGTYEDSTVAEV 175
Qy 143 AGGGCTGCGATACAGCTTACAGTGCATCGATCCCTGCTGCGCAGCTCTGGAAGTC 202
Db 176 EKALEKPEEBSAABEESNDEDEYIPDIVEDVDLNOQVADLNQATTYGWAQDF 235
Qy 203 TGGAAGTACGAATGTTCCCATTAAGGAAGTGTGTGCGCGGCGGCGGCGGCGGCGGCGG 262
Db 236 VRMRKDEBAEALKHAALBEERAMYSGRSRQRREFRERKRLGRKISPPSYARNDSP 295
Qy 263 GCACACTCAGCTGCTTGGCAGGCTTGGGCTTCCAGTACAGGCGGCGGCGGCGGCGGCGGCGG 322
Db 296 TYDPKYSPPSSSSBSRSRSRSPTPGREKITFTTSGSGSEBAAMAAAAAASGVTTGK 355
Qy 323 GCCAGGTGCTGCGAAGGTCCCTCGCAGTCATGAACCAAGGAGGCTTGGGAAACA 382
Db 356 PPAPPGGAPAPGNNAARRRRSSSSSSSSSRSSSSSSSSSSSSSSRRGGGYRSGHA 415
Qy 383 CATCGAAGGCGCATGCTTTCATTTAGTGAAGGTGTGGCTGCGGCGGCGGCGGCGGCGG 442
Db 416 RSRSRWSRSRSRRRSRRSRSGGSRGHRYSRSPARRGGYPRRRRSRSHS 475
Qy 443 AGTCTGAGTACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 502
Db 476 GDRYRRGGRLHHSSSRSSWSLSPSRKSLTRSHSPSPQSRSSRSRSPSPS 535
Qy 503 ATGCAGT 509
Db 536 PAREKLT 542

RESULT 10

US-10-146-731-10
Sequence 10, Application US/10146731
Publication No. US20030128692A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C323
CURRENT APPLICATION NUMBER: US/10/146,731
CURRENT FILING DATE: 2002-05-15


```

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; PRIOR APPLICATION: 2002-05-08
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-10

```

```

Query Match      6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.26; Mismatches 277; Indels 0; Gaps 0;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

```

```

QY 23 AAGACGACGACACTCAGGTGACTACCCCATGTCGAGTGAGGAGCCTCT 82
DB 56 VALAESPVMNMPWQDNTNMI DRFDVRAHLDHPYPLTTISPQESDERKCYER 115
QY 83 GAGGACGAGGCGGACGCGCTCAGGTGCTGACGCGAGGCTCTTGCCATGTCGAC 142
DB 116 YRGLVQNDPAGISBQCLYIYIDELYGGLRPSDEKKLAERKASIGTYEDSTVAEV 175
QY 143 AGGGCTGCTACGACTTACTCAGTGAACAATGAGTCCCTGTCGACGCTCTGGAATC 202
DB 176 EKAKEPBEESAAEBESNSDEVIYPIDIVEVDLQVADLNKQATTYGMADDF 235
QY 203 TGAAGTGAACAATGTTCCATTAAAGAAAGTGTGCGCGCATGCCCCCAACGTT 262
DB 226 VRMLAKDEBAEALIKHAALKEEKAMYSGRSRQRERFRKRLRGRKISPSYARSDP 295
QY 223 GCACACTACCTGCTTTCGAGGCTTGGGGCTTCCAGTCAAGGCTCCATCCAGTACCA 322
DB 296 TYDPKRSPESSSSSRSSRSPPTGREGKITFTISFGSDEBAAAAAAASGVTGK 355
QY 323 GCCCAGGTGCTGACGAAGTCCCTCCAGTCAATGAACCAAGGAGCTTGGAAACA 382
DB 356 PPAPQPGPAPGAPGNAARRSSSSSSASRTSSSSSSSSSRSSRSGGYRSGRHA 415
QY 383 CATCTGAAGGACATGCTTTCATTTAGTGAAGGCTGGGCTGGCTGGGCAAGCCACC 442
DB 416 RSRSSWSRSRSTRSRSSRSGRGRHSGGSRDHRSPARAGGTPRRSRSSHS 475
QY 443 AGGTCTAGTCAAGACGACGAGGACGAAGCTGGTCCCAAGCACTGCCCGCTTGGC 502
DB 476 GDRYRGRGRLRHSSSRSSWSLSPSRSLTRSHSPSPQSRSSRSPSPSPS 535
QY 503 ATGCAGT 509
DB 536 PAREKLT 542

```

```

RESULT 13
US-10-142-885-10
; Sequence 10, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gueney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

```

```

; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; PRIOR APPLICATION: 2002-05-10
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-885-10

```

```

Query Match      6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.26; Mismatches 277; Indels 0; Gaps 0;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

```

```

QY 23 AAGACGACGACACTCAGGTGACTACCCCATGTCGAGTGAGGAGCCTCT 82
DB 56 VALAESPVMNMPWQDNTNMI DRFDVRAHLDHPYPLTTISPQESDERKCYER 115
QY 83 GAGGACGAGGCGGACGCGCTCAGGTGCTGACGCGAGGCTCTTGCCATGTCGAC 142
DB 116 YRGLVQNDPAGISBQCLYIYIDELYGGLRPSDEKKLAERKASIGTYEDSTVAEV 175
QY 143 AGGGCTGCTACGACTTACTCAGTGAACAATGAGTCCCTGTCGACGCTCTGGAATC 202
DB 176 EKAKEPBEESAAEBESNSDEVIYPIDIVEVDLQVADLNKQATTYGMADDF 235
QY 203 TGAAGTGAACAATGTTCCATTAAAGAAAGTGTGCGCGCATGCCCCCAACGTT 262
DB 226 VRMLAKDEBAEALIKHAALKEEKAMYSGRSRQRERFRKRLRGRKISPSYARSDP 295
QY 223 GCACACTACCTGCTTTCGAGGCTTGGGGCTTCCAGTCAAGGCTCCATCCAGTACCA 322
DB 296 TYDPKRSPESSSSSRSSRSPPTGREGKITFTISFGSDEBAAAAAAASGVTGK 355
QY 323 GCCCAGGTGCTGACGAAGTCCCTCCAGTCAATGAACCAAGGAGCTTGGAAACA 382
DB 356 PPAPQPGPAPGAPGNAARRSSSSSSASRTSSSSSSSSSRSSRSGGYRSGRHA 415
QY 383 CATCTGAAGGACATGCTTTCATTTAGTGAAGGCTGGGCTGGCTGGGCAAGCCACC 442
DB 416 RSRSSWSRSRSTRSRSSRSGRGRHSGGSRDHRSPARAGGTPRRSRSSHS 475
QY 443 AGGTCTAGTCAAGACGACGAGGACGAAGCTGGTCCCAAGCACTGCCCGCTTGGC 502
DB 476 GDRYRGRGRLRHSSSRSSWSLSPSRSLTRSHSPSPQSRSSRSPSPSPS 535
QY 503 ATGCAGT 509
DB 536 PAREKLT 542

```

```

RESULT 14
US-10-158-790-10
; Sequence 10, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gueney, Austin L.

```



```

APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumes, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zhenli
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C448
CURRENT APPLICATION NUMBER: US/10/158,790
CURRENT FILING DATE: 2002-05-30
Prior Application removed - See file Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRT
ORGANISM: Homo Sapien
US-10-158-790-10

```

Query Match	6.7%;	Score 37.8;	DB 6;	Length 594;
Best Local Similarity	7.0%;	Pred. No. 0.26;		
Matches	34;	Conservative 176;	Mismatches 277;	Indels 0;
			Gaps	0

Db	56	VALAAESVYNNMPPMGDPTNNMIDLRDVAHLDHIDPYTPPLLTTTSPSEDSERKCNTER	115
QY	83	GAGGACGAGGCCGAGGACCGCTCAGATGGGTACGACGACAGGGTCTTTGCCATGTGGGCAC	142
Db	116	YRGLVQNPFIAGISEQCUIYQIYIDELYGGIQRPSDEKCKLAEKASIGITYEDSVAEV	175
QY	143	AGGGCTCATACACTACTAGTGACATCATGAGTCCCTGGTGGCAGCCTCTGGAGTC	202
Db	176	EKAAPKEPEEESAABEESNSDEVIYPIDIVADVDELNOEQVADLNKATTTYGMDGTF	235
QY	203	TGGAAGTGACATGTTTCCCATYTAAGAAAGTGATGGCCGACCATGCCCCCAACGTT	262
Db	226	VRLMKDNEEALITKAALAEKAMYSRRRRQRREERERKLRKTSPPSYARBDP	295
QY	263	GCACACTCACTGCCTTTGACGAGTTGGGGCTTCAGTCAAGAGTCCCATCCATCAACCA	322
Db	296	TYDPKRSPESSSSRSRSRSPPTQREKXITFTITSFGSDDEAAAAAAGAVTTGK	355
QY	323	GCCCAAGTGGCTGCAGAAAGTCCCTCGCAGTATGAACAACAAGGAGCTTGGGAACA	382
Db	356	PPAPPOGPGAPAGRNASARRRSSSSSSSABRTSSRSSSSSSSRGGGYRSGRA	415
QY	383	CATCGAAGGCGATGGCTTGTGATTTAGTAGAGGGTGGGCTCGGCTGGCAAGGCACC	442
Db	416	RSRSLSWHSRSRSRRTYRSRSRGRHSGGSRDGHRTYRSYPAKRGCTPPRRRSRSRS	475
QY	443	AGGTCTGATGACAGCCAGGACAGGAAAGCTGTCCCGACACTGCCCGCGCTTGGC	502
Db	476	GDRYRGGRGGLHHSSSRSSSMLSPSRSLTTRSHSPSPSGRSRSRSRSGPSPS	535
QY	503	ATGCACT 509	
Db	536	PAREKLT 542	

RESULT 15
US-10-137-871-10
Sequence 10, Application US/10137871
Publication No. US20030207350A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Bersini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Flivartoff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geritsen, Mary E.

```

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C153
CURRENT APPLICATION NUMBER: US/10/137,871
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-871-10

```

Query Match 6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.26;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

QY	23	AAAGAACGAGACACTACCGGTGACTACCCCAATGTCGTGAGAGTGAAGGACCTCT	82
Db	56	VALAAESVVMMPQGGDTNNMIDPFDVAHLDHIDPYIPLLITISPEQESDERKCYER	115
QY	83	GAGCAGAGGCGCAGGCGACGCCGTCAAGTGGGTGACGCGCAGGGGTCTTGCCATGTGGGCAC	142
Db	116	YRGVQVNDFAGISSEQCIFYIYIDELVGGIQRPESEDEKKILAEKASIGIYEDSTVAEV	175
QY	143	AGGGGTGCATACACGCTTACTCAGTGCACATTCGAGTCCCTCGTGCACGCTCTGGAAGTC	202
Db	176	EKAAREPEEESAAEBESNSDEVIDIPIDIVEVDVDELNQVADLNFQATITTYGNADGF	235
QY	203	TGGAAGTAGAGAAATGTTCCCATTAAGGAAGTGTGGCGCGCCATGCCCCCCACGTT	262
Db	236	VRMLRKDEAEALIKHAYALBEEKAMVSGRRSRQRREFREKRLIGRKISPSVYARDSP	295
QY	263	GCACACTACTGCTCTTGACAGGTTGGGGCTTCAGTACACAGGTCCTCATCCAGCTACCA	322
Db	296	TYDPKSPSSSSSESRSRSPPTPGRREKITYFTTSGSSDEBAAAAAAAASVTTGK	355
QY	323	GCCCAAGTGGCTGCAGAAATCCCTCCGACATCATGAACAACGAGGAGGCTTGGGAAACA	382
Db	356	PPAPPGPGPAPAGNNASARRRRSSSSSSSSASRTSSSSSSRSRRGGGYHNSGRHA	415
QY	383	CATCTGAAGGCAATGGCTTTGATTATTTAGTGAAGAGGTGGGGCTGGGGCTGGCAAGCCACC	442
Db	416	RSRGSWSRSRSRSRRYRSRSGRRHSGGGSRDHRSRSPARRGGYGPFRRSRSRSHS	475
QY	443	AGGTCTGAGTCAAGACGAGACGAGAACTGTGTCCTCCAGCACTCCCGCCCTCTCGG	502
Db	476	GDRYRGGCGRLRHSSSSRSRSWLSLPSRSLSLTSHSPSPGSGRSRSRSGSPRS	535
QY	503	ATGCAGT 509	
Db	536	PARCKLT 542	

Search completed: March 25, 2006, 18:14:35
Job time : 682.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 16:11:03 ; Search time 3009.5 Seconds
(without alignments)
10652.835 Million cell updates/sec

Title: SEQ1-69T
Perfect score: 564
Sequence: 1 gcagccctcctcagaaac.....agtgcgtcttcgacttc 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmb1:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_scs:*
11: gb_by:*
12: gb_un:*
13: gb_vl:*
14: gb_hlg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	560.8	99.4	10777	8 HS1013RP	Y12377 H. sapiens F
2	551.4	97.8	150110	14 AC097722	AC097722 Homo sapi
3	551.4	97.8	150214	8 AF006345	AF006345 Homo sapi
4	43	7.6	125020	8 AF429315	AF429315 Homo sapi
5	41.4	7.3	125020	8 AF429315	AF429315 Homo sapi
6	40.2	7.1	320658	14 AC145527	AC145527 Atelelix
7	40	7.1	153129	9 AC163309	AC163309 Mus muscu
8	40	7.1	230041	9 AC111047	AC111047 Mus muscu
9	39.8	7.1	103034	14 AC067978	AC067978 Homo sapi
10	39.8	7.1	111934	8 AF281074	AF281074 Homo sapi
11	39.8	7.1	116039	8 AC007362	AC007362 Homo sapi
12	39.8	7.1	233561	8 AC159874	AC159874 Bos tauru
13	39.6	7.0	119613	8 AC138582	AC138582 Pan trogl
14	39.6	7.0	142879	8 AC000029	AC000029 Homo sapi
15	39.6	7.0	190018	14 AC148588	AC148588 Pan trogl
16	38.8	6.9	184532	9 AC124345	AC124345 Mus muscu
17	38.8	6.9	203106	9 AC021667	AC021667 Mus muscu
18	38.8	6.9	215604	14 AC160979	AC160979 Mus muscu

C 19	38.6	6.8	184533	8	AC093592	AC093592 Homo sapi
C 20	38.6	6.8	192573	8	AC073344	AC073344 Homo sapi
C 21	38.4	6.8	110000	14	CT005255_5	Continuation (6 of
C 22	38.2	6.8	2777	6	AX833454	AX833454 Sequence
C 23	38.2	6.8	2777	8	AK095212	AK095212 Homo sapi
C 24	38.2	6.8	74871	14	AC137095	AC137095 Homo sapi
C 25	38.2	6.8	178778	8	AC132936	AC132936 Homo sapi
C 26	38	6.7	110000	14	LMFLCHR26_2	Continuation (3 of
C 27	38	6.7	121103	8	AP006623	AP006623 Homo sapi
C 28	38	6.7	205854	14	AC138186	AC138186 Homo sapi
C 29	37.8	6.7	113370	8	AC067815	AC067815 Homo sapi
C 30	37.8	6.7	134362	14	AC084853	AC084853 Homo sapi
C 31	37.8	6.7	138129	8	HS102D24	AL021391 Human DNA
C 32	37.8	6.7	186083	9	AC122522	AC122522 Mus muscu
C 33	37.8	6.7	200629	9	AC134463	AC134463 Mus muscu
C 34	37.6	6.7	114694	8	AL355500	AL355500 Human DNA
C 35	37.4	6.6	184866	9	AL611934	AL611934 Mouse DNA
C 36	37.4	6.6	187911	14	AC136867	AC136867 Rattus no
C 37	37.4	6.6	207735	9	AC130279	AC130279 Mus muscu
C 38	37.4	6.6	213590	14	AC154101	AC154101 Mus muscu
C 39	37.4	6.6	223357	14	AC111958	AC111958 Rattus no
C 40	37.4	6.6	240039	14	AC111926	AC111926 Rattus no
C 41	37.2	6.6	21670	8	AF508041	AF508041 Homo sapi
C 42	37.2	6.6	36628	8	AF037222	AF037222 Human DNA
C 43	37.2	6.6	294817	14	AC114626	AC114626 Mus muscu
C 44	37	6.6	1422	4	BT020874	BT020874 Bos tauru
C 45	37	6.6	2500	1	SS055940	SS055940 Streptomyce

ALIGNMENTS

RESULT 1	HS1013RP	10777 bp	DNA	linear	PRI 02-MAR-2000
LOCUS	HS1013RP				
DEFINITION	H. sapiens FGF-3 gene upstream flanking region.				
ACCESSION	Y12377				
VERSION	Y12377.1	GI:1934871			
KEYWORDS	Alu repeat; FGF-3 gene; int-2 gene; L1 repeat; MIR repeat.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	Dienabi, S., Brisson, O., Galdemard, C. and Lavialle, C.			
AUTHORS		Sequence analysis of the transcription control region upstream of			
TITLE		the human FGF-3 gene			
JOURNAL		DNA Seq. 10 (4-5), 317-329 (1999)			
PUBMED		10727086			
REFERENCE		2 (bases 1 to 10777)			
AUTHORS		Brisson, O.			
TITLE		Direct Submission			
JOURNAL		Submitted (04-APR-1997) O. Brisson, Lab. de Genetique Oncologique,			
		CNRSURA1967, Institut Gustave-Roussy, 39 Rue Camille Desmoulins,			
		94805 Villejuif CEDEX, FRANCE			
		The BamHI site at the 3' end of this sequence (nt 10772)			
		corresponds to the BamHI site at the 5' end of the FGF-3/int-2 gene			
		sequence (nt 1; X14445), as ascertained by sequencing through this			
		site.			
		Related sequence: B04369.			
FEATURES		Location/Qualifiers			
Source		1..10777			
		/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
		/chromosome="11"			
		/map="q13"			
		/cissue_type="placenta"			
		/clone_id="lambda D3, Clontech cat#HL10675 lot 1221"			
		/germline			
		489..631			
		/rpt_family="MIR2"			
		repeat_region			

```

repeat_region 632. .945
/tpc_family="Alu"
repeat_region 1084. .1167
/tpc_family="L1MB7"
repeat_region 2605. .2752
/tpc_family="Alu"
repeat_region 3224. .3514
/tpc_family="Alu"
repeat_region 5800. .6089
/tpc_family="Alu"
repeat_region 7003. .7242
/tpc_family="MLTIB"
repeat_region 7417. .7568
/tpc_family="L1MB3"
repeat_region 7715. .7844
/tpc_family="MIR"

```

ORIGIN

```

Query Match 99.4%; Score 560.8; DB 8; Length 10777;
Best Local Similarity 99.6%; Pred. No. 2.2e-122;
Matches 562; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

1 GGAGCCCTGCTCAGAAAACAGAAAGGACGACACACACGTCAGTACACCCCATGTC 60
|
|
|
4945 GGAGCCCTGCTCAGAAAACAGAAAGGACGACACACGTCAGTACACCCCATGTC 5004
|
|
|
61 GCTGAGGTGAGGAGGAGCTCTCTGAGGACAGGACAGGACGCTCAGTGGTGAACGCA 120
|
|
|
5005 GCTGAGGTGAGGAGGAGCTCTCTGAGGACAGGACAGGACGCTCAGTGGTGAACGCA 5064
|
|
|
121 GGGGCTTGGCCATGTTGGGACAGGGGCTGCTACACGCTTACTCAGTGAACATCGATCC 180
|
|
|
5065 GGGGCTTGGCCATGTTGGGACAGGGGCTGCTACACGCTTACTCAGTGAACATCGATCC 5124
|
|
|
181 CTGGTGCCAGCCCTCGAAGTCTGGAAGTGAAGCAATGTTCCATTAAAGAAAGTGTG 240
|
|
|
5125 CTGGTGCCAGCCCTCGAAGTCTGGAAGTGAAGCAATGTTCCATTAAAGAAAGTGTG 5184
|
|
|
241 GCCGGCCATCCGCCCAAGTTCACACTCTGCTTTCAGAGGTTGGGGCTTCCAGTC 300
|
|
|
5185 GCCGGCCATCCGCCCAAGTTCACACTCTGCTTTCAGAGGTTGGGGCTTCCAGTC 5244
|
|
|
301 AAGAGGTCCCATTCACATGACAGCCAGGTGCTGCAAGAGTCCCTGCGCATGTA 360
|
|
|
5245 AAGAGGTCCCATTCACATGACAGCCAGGTGCTGCAAGAGTCCCTGCGCATGTA 5304
|
|
|
361 CCAAGGAGGAGGCTTGGGAAACCATCTGAAGGAGGAGTGGTTGATTTATGAGAGGTCG 420
|
|
|
5305 CCAAGGAGGAGGCTTGGGAAACCATCTGAAGGAGGAGTGGTTGATTTATGAGAGGTCG 5364
|
|
|
421 GCGTGGGCTGGGCAAGGACCAAGTCTGAGTCAGAGCCAGAGGACGAGAGCTGGTCCC 480
|
|
|
5365 GCGTGGGCTGGGCAAGGACCAAGTCTGAGTCAGAGCCAGAGGACGAGAGCTGGTCCC 5424
|
|
|
481 AGCACTGCCCCGCGCTCTGCGATGATCTCTCTGGCCACCTGAGAAAGCCTGTAGAG 540
|
|
|
5425 AGCACTGCCCCGCGCTCTGCGATGATCTCTCTGGCCACCTGAGAAAGCCTGTAGAG 5484
|
|
|
541 AGGCACTGGCGCTCTTCCGACATTC 564
|
|
|
5485 AGGCACTGGCGCTCTTCCGACATTC 5508
|
|
|

```

```

RESULT 2
AC097722/c AC097722 150110 bp DNA 1linear HTG 24-JAN-2002

```

```

DEFINITION Homo sapiens chromosome UNK clone RP11-186D19, WORKING DRAFT
SEQUENCE 9 unordered pieces.

```

```

ACCESSION AC097722

```

```

VERSION AC097722.2 GI:19308824

```

```

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS 1 (bases 1 to 150110)
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
JOURNAL 2 (bases 1 to 150110)
AUTHORS Waterston,R.H.
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jan 24, 2002 this sequence version replaced gi:16306472.

```

COMMENT

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: H.NH0186D19

```

```

----- Summary Statistics -----
Sequencing vector: M13; 54%
Sequencing vector: plasmid; 46%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 147449 bases at least Q40
Consensus quality: 148263 bases at least Q30
Consensus quality: 148688 bases at least Q20
Insert size: 161000; agarose-fp
Insert size: 150429; sum-of-contigs
Quality coverage: 8.73 in Q20 bases; agarose-fp
Quality coverage: 7.35 in Q20 bases; sum-of-contigs

```

```

* NOTE: This is a "working draft" sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 3944: contig of 3944 bp in length
* 3945 4044: gap of unknown length
* 4045 9798: contig of 5754 bp in length
* 9799 9898: gap of unknown length
* 9899 14019: contig of 4121 bp in length
* 14020 14119: gap of unknown length
* 14120 19392: contig of 5273 bp in length
* 19393 19492: gap of unknown length
* 19493 24872: contig of 5380 bp in length
* 24873 24972: gap of unknown length
* 24973 43958: contig of 18986 bp in length
* 43959 44058: gap of unknown length
* 44059 68992: contig of 24934 bp in length
* 68993 104669: gap of unknown length
* 104670 104769: gap of unknown length
* 104770 150110: contig of 45341 bp in length.

```

FEATURES

```

source 1..150110
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="UNK"
/clone="RP11-186D19"
misc_feature 1..3944
/note="assembly_name:Contig25"
gap 3945..4044
/estimated_length=unknown

```

```
misc_feature 4045. .9798 /note="assembly_name:Contig25"  
gap 9799. .9898  
misc_feature /estimated_length=unknown  
9899. .14019  
gap /note="assembly_name:Contig27"  
14020. .14119  
misc_feature /estimated_length=unknown  
14120. .19392  
gap /note="assembly_name:Contig28"  
19393. .19492  
misc_feature /estimated_length=unknown  
19493. .24872  
gap /note="assembly_name:Contig29"  
24873. .24972  
misc_feature /estimated_length=unknown  
24973. .43958  
gap /note="assembly_name:Contig30"  
43959. .44058  
misc_feature /estimated_length=unknown  
44059. .68992  
gap /note="assembly_name:Contig31"  
68993. .69092  
misc_feature /estimated_length=unknown  
69093. .104669  
gap /note="assembly_name:Contig32"  
clone_end:T7  
vector_side:left  
104670. .104769  
misc_feature /estimated_length=unknown  
104770. .150110  
gap /note="assembly_name:Contig33"  
clone_end:SP6  
vector_side:right
```

Query Match	97.8%	Score 551.4;	DB 14;	Length 150110;
Best Local Similarity	99.6%	Pred. NO. 2.7e-120;		
Matches 563; Conservative	0;	Mismatches 1;	Indels 1;	Gaps 1

QY	1	GCAGCCCTGCGCTCAGAAAAACGAAAGACCGAGCAACTCAACGGGTGATCAACCCCAATG	60
Db	50478	GACGCCCTGCTTCAGAAAAACGAAAGACCGAGCAACTCAACGGGTGATCAACCCCAATG	50411
QY	61	GCTGAGGTGAGGAGGAGCCTCTCTGAGCGAGGGCCAGGGCAGGCCGTCAAGTGGGTGACGGCA	120
Db	50418	GCTGAGGCGAGGGAGCCTCTCTGAGGCGAGGGCCAGGGCAGGCCGTCAAGTGGGTGACGGCA	50355
QY	121	GGGCTCTTGGCCATGGTGGGGCAACGGGGCTGCAATCACTTAATCTAGTGAACAATCGAGTCC	180
Db	50358	GGGCTCTTGGCCATGGTGGGGCAACGGGGCTGCAATCACTTAATCTAGTGAACAATCGAGTCC	50295
QY	181	CTGGTGCCAGCCTCTGGAAGTCTGGAAGTGAAGCATGTTTCCCATTAAGGAAAGTGTTG	240
Db	50298	CTGGTGCCAGCCTCTGGAAGTCTGGAAGTGAAGCATGTTTCCCATTAAGGAAAGTGTTG	50233
QY	241	GCCGGCCATGCCCCCAACGTTTGCAACTCACTGCTCTTTGACAGGGTGGGGCTTCCAGTCC	300
Db	50238	GCCGGCCATGCCCCCAACGTTTGCAACTCACTGCTCTTTGACAGGGTGGGGCTTCCAGTCC	50175
QY	301	ACAGGGTCCCATCCACGTACCAAGGCCCAAGTGGCTGCAAGAAAGTTCCTCGCAGTCAAGAA	360
Db	50178	ACAGGGTCCCATCCACGTACCAAGGCCCAAGTGGCTGCAAGAAAGTTCCTCGCAGTCAAGAA	50115
QY	361	CCAAAGGAGGCTTGGGAAACCAATCTGAAGGCGATGGCTTTGATTTAATGTAAGAGGGTGG	420
Db	50118	CCAAAGGAGGCTTGGGAAACCAATCTGAAGGCGCGATGGCTTTGATTTAATGTAAGAGGGTGG	50055
QY	421	GGCTGGGCTTGGGCAAGGCAACAGGTCTGAATGAGAGCCAGAGGCAAGGCAAGGCTGGTCCCC	480
Db	50058	GGCTGGGCTTGGGCAAGGCAACAGGTCTGAATGAGAGCCAGAGGCAAGGCAAGGCTGGTCCCC	49995
QY	481	A-GCATTGCCCCCGCCCTTGCGAATGCAATGCTCTTGAGCAACTCTGAAGAACGCTGTGA	539

Db 49998 AGGACACGCCCGCGCTCTGCATGAGTCTCTTGGSCACCTGAGAAACAGCTGTAGA 49999

Oy 540 GAGGCAGTGGCGTCTTTCGACTTC 564

Db 49938 GAGGCAGTGGCGTCTTTCGACTTC 49914

[illegible]

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE	AUTHORS	TITLE
1	Hattori, M., Toyoda, A., Taylor, T. D., Kuroki, Y., Fujiyama, A., Totoki, Y., and Sakaki, Y.	Homo sapiens genomic DNA

JOURNAL Published Only in Dataase (2003)
REFERENCE 2 (bases 1 to 150214)
AUTHORS Hattori, M., Toyoda, A., Taylor, T. D., Kuroki, Y., Fujiyama, A., Totoki, Y. and Sakaki, Y.

TITLE Direct Submission
JOURNAL Submitted (30-APR-2003) Masahira Hattori, The Institute of Physical

JOURNAL
Submitted (30-APR-2003) Masahira Hatori The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail: hatori@gscc.riken.jp, URL: <http://bgp.gsc.riken.jp/>,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT
On Aug 5, 2005 this sequence version replaced gi:48290861.

COMMENT	On Aug 5, 2005 this sequence version replaced gi:48290861
FEATURES	location/Qualifiers
source	1. .150214

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-186D19"

```

Query Match	97.8%	Score 551.4	DB 8	Length 150214
Best Local Similarly	99.6%	Pred. No. 2.7e-120		
Matches 563	0	Mismatches 1	Indels 1	Gaps 1

Oy	GCAGCCCTGCTCAGAAAAACAAGGACGAGACACTCAGGTATCTCACCCCATGTG	60
Db	85889 GCAGCCCTGCTCAGAAAAACAAGGACGAGACACTCAGGTATCTCACCCCATGTG	85830
Oy	61 GCTGAGGTAGGGAGCCTTCTTAGGGCAGAGGCCACGCGTCAAGTGGGTGACGGCA	120
Db	85829 GCTGAGGTAGGGAGCCTTCTTAGGGCAGAGGCCACGCGTCAAGTGGGTGACGGCA	85770
Oy	121 GGGGCTCTTGGCATGGTGGGCAACAGGGGCTGCAATCAGCTTACTCAGTGCACATGCAATGCC	180
Db	85769 GGGGCTCTTGGCATGGTGGGCAACAGGGGCTGCAATCAGCTTACTCAGTGCACATGCAATGCC	85710
Oy	181 CTGGTGCACAGCCTCTTGAAAGTCTGAAAGTGAGCAATGTTTCCCATTTAAGAAAGTGTGTG	240
Db	85709 CTGGTGCACAGCCTCTTGAAAGTCTGAAAGTGAGCAATGTTTCCCATTTAAGAAAGTGTGTG	85650
Oy	241 GCCGGGCATGCCCCCAAGGTTGCAACATCATGCGCTTTGGCAGGGTGGGGCTTCACATGC	300
Db	85849 GCCGGGCATGCCCCCAAGGTTGCAACATCATGCGCTTTGGCAGGGTGGGGCTTCACATGC	85590
Oy	301 ACAAGGTCCTCATCAGTACACAGCCACAGGTGCTGACAAAGGTCTCTCGCAATCATGAAA	360
Db	85589 ACAAGGTCCTCATCAGTACACAGCCACAGGTGCTGACAAAGGTCTCTCGCAATCATGAAA	85530

```
QY 361 CCAAGGAGAGGCTTGGAAAACAATCTGAAAGGCATGCTTGAATTAGTAGAGAGGGTGG 420
Db 8529 CCAAGGAGAGGCTTGGAAAACAATCTGAAAGGCATGCTTGAATTAGTAGAGAGGGTGG 85470
QY 421 GCGTGGGCTGGGCAAGGCCACAGGCTCTGATGACAGGACGAGGAGGAAGCTGTCCCC 480
Db 85469 GCGTGGGCTGGGCAAGGCCACAGGCTCTGATGACAGGACGAGGAGGAAGCTGTCCCC 85410
QY 481 A-GCACTGCCCCCGCTCTGCGATGCACTCTCTGCGCCACCTGAGAACAGCTGTAGA 539
Db 85409 AGGCACTGCCCCCGCTCTGCGATGCACTCTCTGCGCCACCTGAGAACAGCTGTAGA 85350
QY 540 GAGGCAGTGGCGCTCTTTCGGAAGCTTTC 564
Db 85349 GAGGCAGTGGCGCTCTTTCGGAAGCTTTC 85325

RESULT 4
AF429315/c 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Porter,N.T., Ross,C.A. and Margolis,R.L.
TITLE A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
source 1..125020
/mol_type="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
disease-like 2 (HDL2)"
complement(35581..35746)
repeat_region
rpt_type=tandem
rpt_unit="ctg"
complement(<36507..>36887)
/gene="JPH3"
/note="synonym: JPH3"
complement(<36507..>36887)
/gene="JPH3"
/product="junctophilin 3"
complement(<36507..>36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AAL40941.1"
/db_xref="GI:17646245"
/translation="MSSGGRFNFDDGGSGYCGMWDGKAKHGVCCTPKGQGGYTSMS
HGFVLYTWPDSNTYGGTAAQGRHGLGSKKRWYKGMWTHGFGRGVRECAQ
NGAKYEGTWSNGLDQYGTETYSDG"
```

```
Query Match 7.6%; Score 43; DB 8; Length 125020;
Best Local Similarity 12.0%; Pred. No. 6.2;
Matches 64; Conservative 221; Mismatches 245; Indels 2; Gaps 1;

QY 23 AAGACGAGACACACTACCGGTGACTACCCCCCATGTGGCTGAGAGTAGAGGAGGCTCT 82
Db 17541 MRSSKSSWSMSMARRSSWCMGAGYRRSKRSAGCAWMSKGRSTGMKACSKTKGS 17482
QY 83 GAGGACAGGCGCAGGACAGCCCTCAGGTGAGGAGCGGAGGAGTCTTGCATGTGTGGAC 142
Db 17481 YGSTRSMKKKKKXGYSKYSRKGKKKTKTCYKMKYKTKKMSCMWYKMSNGYKXKXC 17422
QY 143 AAGGCTGCATATACAGCTTACTACTAGTACATATGATCCCTGAGCTCCAGCTTGAAGTC 202
Db 17421 MKKKGCTYRGMSSKSKSGYKSMGRGSSYSTSCWKSQWYSMMKCMKYMKKRRBMR 17362
QY 203 TGAAGTACGCAATGTTCCATTAAAGAAAGTGTGTGCGCCGACATCCCCCAAGCTT 262
Db 17361 GMSKSKMGRGYAGRGYSSSSWMSSTRKRSKCYSKYKRGKKGKMGWGMKRGSKYWSM 17302
QY 263 GCACACTGACTGCTTTCGAGGTTGGGCTTCCAGTACAGAGGCTCCATCAGTACCA 322
Db 17301 KMKRSMWSKCTYKYSKGRSKGRSKGRSTKSKAKSSMRAGSKCTYGSYSWNNRNN 17242
QY 323 GCCCAGGTGGCTGCAAGAGT--CCCTGCAATCATGAACAAGAGGCTTGGAAAC 380
Db 17241 RMKGTGCVMYRRBRAMNNGNAAAAGCTTCCCCANTNGGGGAAAAGGCGSASRASCY 17182
QY 381 CACATCTGAAGGACATGCTTGTATTAGTAGAGAGGTGGGCTGGCTGGGCAAGGCCA 440
Db 17181 KGRSSSKSCYSGTTRRRKMKCKRRSGKMGMTRGSGGKTSYANGSGCYCWW 17122
QY 441 CCAGTCTGAGTACAGAGCAGGACGAGCAAGACTGTGCTCCCACTGCGCCCGCTGTG 500
Db 17121 GRKGRKMSKMKMYKYSYRKMKMTCKMTKMTCTMTAMYGRSMCMCKSCCGCTSMG 17062
QY 501 CGATGACTCTCTCTGGCCACTGAGAACACAGCTGTAGAGAGGAGGAGTGGCGT 552
Db 17061 MSSYSKYSKMGMSKMYMRSYSKRSKAMRSKRMGMGTGGRYGGGRSY 17010

RESULT 5
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Porter,N.T., Ross,C.A. and Margolis,R.L.
TITLE A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
source 1..125020
/mol_type="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
```

/map="16q24.3; between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
disease-like 2 (HDL2)"
repeat_region complement(35581..35746)
/rpt_type=tandem
/rpt_unit="ctg"
gene /gene="JPH3"
complement(<36507..36887)
/note="synonym: JP3"
complement(<36507..36887)
/gene="JPH3"
/product="junctional protein 3"
complement(<36507..36887)
/note="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
CDS /codon_start=1
/product="junctional protein 3"
/protein_id="AA140941.1"
/db_xref="GI:17646245"
/translation="MSGGRNFDDGSGYCGMEDGKAHGHVCTGPKGQEGTYSMS
HGFEVLYTTPSGTGGTGAAGCRHIGLESKQKMYKGEWTHGFKRGVRECAH
NGAKYEGTMSGLDGYETERTSDG"

ORIGIN

Query Match 7.3%; Score 41.4; DB 8; Length 125020;
Best Local Similarity 11.5%; Pred. No.15;
Matches 45; Conservative 169; Mismatches 176; Indels 0; Gaps 0;

OY 32 GCACACTACGCTGACTACACCCCATGTGGCTGAGTGAAGGAGCCTCTGAGCAGAGG 91
DB 51746 KSAVSCMWSKXKCKSSYCAKSGCYKXKYMWSYTSGRSMWYTSYSCMYMSMMWSY 51805
OY 92 CCAGGAGCAGCCTCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 151
DB 51806 MSWSKSGYTSKCMKCMKCMKCMKCMKCMKCMKCMKCMKCMKCMKCMKCMKCMK 51865
OY 152 ATACAGCTTACTCACTGACCAATCGAGTCCCTGGTCCAGCCTCTGGAAGTCTGGAAGA 211
DB 51866 RRSCKMSKTKGTCCKKSKSMKCATTKMKSCAKMSKSCYCGMKCRASCMGCKYKMSKT 51925
OY 212 GCAATGTTTCCCATTAAGAAAGTGTGGCCGGCCATGCCCCCAAGTTCACACTCA 271
DB 51926 CMMSYTTSCYCYGCMKMKKWKRRSSYKSYCKKSMTKYSWRGMSCYRGRSYVS 51985
OY 272 CTGCTTTGACAGGTTGGGCTTCAGTACAGAGGTCCTCCATCCAGTACAGCCAGGTG 331
DB 51986 RYRMKMKCMGWTTRKSKSKSSNSSMGMMWGKTSYMCASSYASCKRGRSATSWGCC 52045
OY 332 GCTGCAAGAGTCCCTCGACATCATGAACCAAGGAGGAGGAGGAGGAGGAGGAGG 391
DB 52046 YMSYRSASCMGSGTSMRSYSTRMSCHMSCHMSCHMSCHMSCHMSCHMSCHMSCH 52105
OY 392 GGCATGCTTGTGATTAGTAGAGAGGTGG 421
DB 52106 RGSWMKSGMTKSKSGTGKGMWGRSGWSKS 52135

RESULT 6
AC145527/c 320658 bp DNA linear HTG 19-JUL-2003
LOCUS AC145527

DEFINITION Atelelex albiventris clone LB4-283G23, WORKING DRAFT SEQUENCE, 21
unordered pieces.

ACCESSION AC145527
VERSION AC145527.1 GI:3296768
KEYWORDS HTG, PHASE1, HTGS DRAFT.
SOURCE Atelelex albiventris (middle-African hedgehog)
ORGANISM Atelelex albiventris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Insectivora; Etriacidae;
Etriacineae; Atelelex.
REFERENCE 1 (bases 1 to 320658)
AUTHORS Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,

TITLE Peng,Z., Malinov,I. and Rubin,E.M.
JOURNAL Direct Submission
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 320658)
Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
Peng,Z., Malinov,I. and Rubin,E.M.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
Draft Sequence Produced by Berkeley PGA
Web site: <http://pga.lbl.gov>
Center Code: PGABERK
Center Project Name: E024-283G23
Bac Clone Name: LB4-283G23

Additional information on comparative analysis and ordering are
available at:
http://pga.lbl.gov/cgi-bin/search_cvs.cgi?type=na&value=SRBFP1
Funding agent: Programs for Genomic Applications (NHLBI)
if library name is LBI to LB4, please see website
for a description: <http://www.gsdl.lbl.gov/cheng/BAC.html>
These libraries are available through the BACPAC Resources Center:
<http://www.chori.org/bacpac/libraries.htm> as LBML-1 to LBML-4.
Summary Statistics:
Sequencing vector: Plasmid; pUC18
Chemistry: Dye-terminator Big Dye
Assembly program: Phrap version 0.990329.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2325: contig of 2325 bp in length
2326 2425: gap of unknown length
2426 4760: contig of 2335 bp in length
4761 4860: gap of unknown length
4861 7844: contig of 2984 bp in length
7845 7944: gap of unknown length
7944 10691: contig of 2747 bp in length
10692 10791: gap of unknown length
10792 15758: contig of 4967 bp in length
15759 15858: gap of unknown length
15859 18758: contig of 2900 bp in length
18759 18858: gap of unknown length
18859 23515: contig of 4657 bp in length
23516 29133: gap of unknown length
29134 29133: contig of 5518 bp in length
29134 29234: gap of unknown length
29234 33870: contig of 4637 bp in length
33871 33970: gap of unknown length
33971 33971: contig of 5546 bp in length
33971 39616: gap of unknown length
39617 45912: contig of 6296 bp in length
45913 46012: gap of unknown length
46013 52370: contig of 6358 bp in length
52371 52470: gap of unknown length
52471 64704: contig of 12234 bp in length
64705 76264: gap of unknown length
76265 76264: contig of 11460 bp in length
76265 76365: gap of unknown length
76365 94758: contig of 18394 bp in length
94759 94859: gap of unknown length
94859 118713: contig of 23855 bp in length
118714 118813: gap of unknown length
118814 148021: contig of 29208 bp in length
148022 148121: gap of unknown length
148122 178939: contig of 30818 bp in length
178940 211754: gap of unknown length
211755 211754: contig of 32715 bp in length
211755 211854: gap of unknown length

FEATURES	
SOURCE	Location/Qualifiers
	1. .153129
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="3"
	/map="3"
	/clone="RP24-11708"
	/clone_1lb="RP21-24 Male Mouse BAC"
repeat_region	2843. .2893
repeat_region	/rpt_family="AT_rich"
	complement(12955. .3180)
repeat_region	/rpt_family="B4"
	544. .5470
repeat_region	/rpt_family="GAA"n"
	5748. .5876
repeat_region	/rpt_family="TCTA"n"
	6984. .7110
repeat_region	/rpt_family="MLT1C"
	7462. .7801
repeat_region	/rpt_family="ORR1D"
	complement(19528. .9910)
repeat_region	/rpt_family="MT2B"
	9937. .10085
repeat_region	/rpt_family="MERSB"
	10825. .10982
repeat_region	/rpt_family="TG"n"
	complement(11051. .11261)
repeat_region	/rpt_family="B4"
	12012. .12033
repeat_region	/rpt_family="(A) n"
	12907. .13123
repeat_region	/rpt_family="MLT1C"
	13433. .13812
repeat_region	/rpt_family="MTE"
	complement(14278. .14503)
repeat_region	/rpt_family="L1"
	complement(14417. .14746)
repeat_region	/rpt_family="L1"
	14745. .15433
repeat_region	/rpt_family="L1_MM"
	complement(15710. .15820)
repeat_region	/rpt_family="L1MA4A"
	complement(15523. .16050)
repeat_region	/rpt_family="L1MA4A"
	16057. .16122
repeat_region	/rpt_family="(TG) n"
	complement(16123. .16262)
repeat_region	/rpt_family="B3A"
	16576. .16671
repeat_region	/rpt_family="(TG) n"
	16676. .16856
repeat_region	/rpt_family="(TATATG) n"
	16862. .16925
repeat_region	/rpt_family="(TATG) n"
	16926. .17091
repeat_region	/rpt_family="(TATATG) n"
	complement(17211. .17610)
repeat_region	/rpt_family="MLT1A0"
	complement(17617. .19453)
repeat_region	/rpt_family="L1_MM"
	18133. .18163
unbure	/note="single clone coverage"
	19454. .20704
repeat_region	/rpt_family="L1"
	20758. .20824
repeat_region	/rpt_family="MIR"
	complement(20852. .21030)
repeat_region	/rpt_family="MIR"
	21915. .21942
repeat_region	/rpt_family="(TTTTTG) n"
	complement(21943. .22157)

repeat_region	/rptc family="B3" 22301. .22397 /rptc family="l2" complement(22732. .22787) /rptc family="l1" 22786. .23404 /rptc family="l1_MM" complement(23405. .23879) /rptc family="MERVL_LTR" 23880. .24023 /rptc family="l1_MM" 24051. .24205 /rptc family=" (TCTCC) n" 25614. .25638 /rptc family="AT_rich" 26685. .26707 /rptc family=" (TTGC) n" 28871. .28954 /rptc family=" (TA) n" complement(29240. .29627) /rptc family="TAPEX_LTR" complement(29660. .29993) /rptc family="ORR1D" 30197. .30223 /rptc family=" (TTTC) n" complement(30226. .30374) /rptc family="B1_MM" 30733. .30785 /rptc family=" (TTTC) n" complement(30891. .31095) /rptc family="AIR" complement(30926. .31037) /note="single clone coverage" 31116. .31151 /rptc family=" (GGGGA) n" 31130. .31168 /note="single clone coverage" 31345. .31410 /rptc family=" (TG) n" complement(31417. .31571) /rptc family="RSINB1" complement(33166. .33865) /rptc family="RMERGA" 38071. .38261 /rptc family="B3A" 38280. .38428 /rptc family="B3A" 38402. .38558 /rptc family=" (CA) n" 39152. .39203 /rptc family=" (TTA) n"
---------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	7.1k	Score 40	DB 9	Length 153129
Query Match	Best Local Similarity	52.4%	Pred. No. 31	
Matches	88	Conservative	0	Mismatches 80, Indels 0, Gaps 0
Qy	301	ACAAGGTCCTCCATCCAGTACACGACCCAGGTGGCTGCAGAAAGTCCCTCCAGTCATGAAA	360	
Db	76180	ACAAATGATGCTGCTGGAGCCAGGCGAGAGGCAAGGGCGAGGGCCAGTGCAGAGCCAAAG	761211	
Qy	361	CCAAAGGAGGCTTGGGAAAACCATCTGAAAGGCATGGCTTTGATTACTGAGAGAGGTGG	420	
Db	76120	GACAGGAGCCAGGAGTGTGAAGCTTAGGTTAGAGAGCAGGCGGGGCCAGACAGAAAGCCAGG	760615	
Qy	421	GGCTGGGCTGGGCAAGGCCACCAAGCTGTAGTTCAGAGCCAGAGGCGCAGG	468	
Db	76060	GGCGGGGCCAGGCGCAGGAGGCAAGGCGCAGGCGCAGGCGCAGGCGCAGG	76013	
RESULT 8				
LOCUS	AC111047	230041 bp	DNA	linear
DEFINITION	Mus musculus chromosome 3, clone RP23-17EH24, complete sequence.			
LOCUS	AC111047	230041 bp	DNA	linear
DEFINITION	Mus musculus chromosome 3, clone RP23-17EH24, complete sequence.			

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AC111047.9	GI:37620322	HTG.	Mus musculus (house mouse)	Mus musculus
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				</

FEATURES	Location/Qualifiers
source	1..231041 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="3" /map="3" /clone="RP23-176H24" /clone_1fb="RPCI-23 Female Mouse BAC" 1..21291 /note="wgs end extension clone_end:Sp6"
misc_feature	repeat_region 1144..1290 /rpt_family="B1_MM" repeat_region 2418..2629 /rpt_family="L2" repeat_region 2641..2696 /rpt_family="MER91A" repeat_region 3023..3043 /rpt_family="(TCTG)n" repeat_region 3086..3126 /rpt_family="(CA)n" repeat_region 3126..3176 /rpt_family="(GA)n" repeat_region 3293..3449 /rpt_family="L2" repeat_region 3572..3576 /note="<30 qual SNGU region" complement(5063..5383) /note="single clone coverage" unsure 5464..5470 /note="<30 qual SNGU region" unsure 5498..5503 /note="<30 qual SNGU region" unsure 5539..5544 /note="<30 qual SNGU region" unsure 5564..5568 /note="<30 qual SNGU region" repeat_region /rpt_family="Lx2B" 6272..6332 /rpt_family="(GA)n" repeat_region 7692..7878 /rpt_family="ORR1B1" complement(8072..8141) /rpt_family="MIR" 8906..9058 /rpt_family="B4A" repeat_region 9622..9758 /rpt_family="(TCC)n" complement(10521..10594) /rpt_family="zapdog" repeat_region complement(12702..12858) /rpt_family="B3" complement(12871..13230) /rpt_family="MTD" 13255..13275 /rpt_family="AT_rich" repeat_region 13286..14525 /rpt_family="Lx4" 14031..14049 /note="single clone coverage" 14535..14580 /rpt_family="(CAAT)n" complement(14607..14691) /rpt_family="Lx9"


```
repeat_region 14812. .14846 /rpt_family="CAAAA)n"
repeat_region 15205. .15348 /rpt_family="B3A"
repeat_region 15349. .15391 /rpt_family="CA)n"
repeat_region 15392. .15437 /rpt_family="B3A"
repeat_region complement(15846. .15975) /rpt_family="MTD"
repeat_region 16110. .16376 /rpt_family="MTC"
repeat_region 18724. .18891 /rpt_family="L1MB8"
repeat_region 19356. .19561 /rpt_family="B3"
repeat_region 19961. .19988 /rpt_family="AT_r1ch"
repeat_region 19990. .20035 /rpt_family="TCTA)n"
repeat_region 20124. .20180 /rpt_family="AT_r1ch"
repeat_region complement(20343. .20479) /rpt_family="B1_MM"
repeat_region complement(20694. .21089) /rpt_family="MTA"
misc_feature 21297. .21297 /note="clone boundary"
clone_end:sp6
site:ECOR1
repeat_region 21395. .21487 /rpt_family="CAGGG)n"
repeat_region 21517. .21672 /rpt_family="L1MB8"
repeat_region complement(27914. .28115) /rpt_family="URR1A"
repeat_region 30025. .30065 /rpt_family="(TG)n"
repeat_region complement(30071. .30214) /rpt_family="KSINE1"
repeat_region 30216. .30311 /rpt_family="CT_r1ch"
repeat_region complement(30413. .30611) /rpt_family="B4"
repeat_region 31033. .31131 /rpt_family="(CCA)n"
repeat_region 32917. .33316 /rpt_family="(TG)n"
repeat_region 33334. .33515 /rpt_family="(TG)n"
repeat_region 33712. .33732 /rpt_family="AT_r1ch"
```

Query Match	7.1*	Score 40;	DB 9;	Length 230041;
Best Local Similarity	52.4*	Pred. No. 29;		
Matches	89;	Conservative	0;	Mismatches 80; Indels 0; Gaps 0;
Qy	301	ACAGGATCCATCCACGTACCAAGCCACAGGCTGGCGAAGAGTCCCTCGACATCATGAAA	360	
Db	21304	ACAAAGTAGCTGCTGGAGCCAGGGCAGAGAGCCAGGGGCGCATGTGCAGAGCCAA	21365	
Qy	361	CCAAAGGAGGCTTTGGGAAACCATCTGAAGGGCATGGCTTTGATTAGTAGAGGGTGG	420	
Db	21364	GACACGAGCCACGAGGGGTGAGCTAGGTATGAGAGGCGAGGGGCCAGAGCCAAAGGCAGG	21423	
Qy	421	GGCTGGGCTGGGGCAAGGCCACAGAGTGTGAATCCAGACCAAGAGGACAG	468	
Db	21424	GGGGGGGCGAGGCCAGAGGAGGAGGGGCGAGGCCAGGCCAGAGGACAG	21471	

RESULT 9
AC067978
AC067978

AC067978 103034 bp DNA linear HTG 03-JUL-2000

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
Homo sapiens chromosome 2 clone CTD-2350X7 map 2, WORKING DRAFT SEQUENCE. 15 unordered pieces.	AC067978	AC067978.2	GI:8901271	HTG, HTGS PHASE1, HTGS_DRAFT.	Homo sapiens (human)	1 (bases 1 to 103034)	Bitren,B., Linton,L., Nusbaum,C. and Lander,E.
						2 (bases 1 to 103034)	Bitren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	
						Unpublished	
						1 (bases 1 to 103034)	
						2 (bases 1 to 103034)	

TITLE	Direct Submission
JOURNAL	Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Jul 3, 2000 this sequence version replaced gi:7656720.

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: 19518
Center clone name: 2350 X 7
-----
Summary Statistics
Sequencing vector: M13; M77815, 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 95545 bases at least Q40
Consensus quality: 99446 bases at least Q30
Consensus quality: 100858 bases at least Q20
Insert size: 98000; agarose-fp
Insert size: 101634; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1585: contig of 1585 bp in length
1586
1685: gap of 100 bp

```

*	1686	3150: contig of 1465 bp in length
*	3151	3250: gap of 100 bp
*	3251	4742: contig of 1492 bp in length
*	4743	4842: gap of 100 bp
*	4843	6347: contig of 1505 bp in length
*	6348	6447: gap of 100 bp
*	6448	8790: contig of 2343 bp in length
*	8791	8890: gap of 100 bp
*	8891	11684: contig of 2794 bp in length
*	11685	11784: gap of 100 bp
*	11785	15905: contig of 4121 bp in length
*	15906	16005: gap of 100 bp
*	16006	21395: contig of 5390 bp in length
*	21396	21495: gap of 100 bp
*	21496	26064: contig of 4569 bp in length
*	26065	26164: gap of 100 bp
*	26165	34912: contig of 8748 bp in length
*	34913	35012: gap of 100 bp
*	35013	46466: contig of 11454 bp in length
*	46467	46566: gap of 100 bp
*	46567	56982: contig of 10416 bp in length
*	56983	57082: gap of 100 bp
*	57083	71308: contig of 14226 bp in length
*	71309	71408: gap of 100 bp
*	71409	86559: contig of 13151 bp in length
*	86560	86659: gap of 100 bp
*	86660	103034: contig of 16375 bp in length

FEATURES	Location/Qualifiers
source	1. .103034

	gap	/note="assembly_fragment"	34913.	.35012		
		/estimated_length=100				
	misc_feature	35013..46466				
		/note="assembly_fragment"				
	gap	46467..46566				
		/estimated_length=100				
	misc_feature	46567..56982				
		/note="assembly_fragment"				
	gap	56983..57082				
		/estimated_length=100				
	misc_feature	57083..71308				
		/note="assembly_fragment"				
	gap	71309..71408				
		/estimated_length=100				
	misc_feature	71409..86559				
		/note="assembly_fragment"				
	gap	86560..86659				
		/estimated_length=100				
	misc_feature	86660..103034				
		/note="assembly_fragment"				
ORIGIN						
Query Match		7.1%; Score 39.8; DB 14;	Length	103034;		
Best Local Similarity	64.8%;	Pred. No. 36;				
Matches	59;	Conservative	0;	Mismatches	32;	
			Indels	0;	Gaps	0;
Oy	4	GCCTGCTCAGAAAACAGAAAGACGACACTCACGGTGACTCACCCCCATGTGGCT	63			

Db	92987	GCCCTCCCTCGAAGAACGAGGACCCCTCGGCGCTACCCCTCATTTGGAA	93046
Qy	64	GGAGGTGAGGAGACCTTCCTGAGCGCAGGCCA	94
Db	93047	TGATGAGAGGTGTTCCCAAGCGCTGGCCA	93077

```

RESULT 10
AF281074/c AF281074 111934 bp DNA 1linear PRI 21-DEC-2000
LOCUS Homo sapiens neuropilin 2 (NRP2) gene, complete cde, alternatively
DEFINITION spliced.
ACCESSION AF281074
VERSION AF281074.1 GI:11934947
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 111934)
Rossignol,M., Gagnon,M.L. and Klasebrun,M.
Genomic organization of human neuropilin-1 and neuropilin-2 genes:
identification and distribution of splice variants and soluble
isoforms
JOURNAL Genomics 70 (2), 211-222 (2000)
PUBMED 11112349
2 (bases 1 to 111934)
Rossignol,M., Gagnon,M.L. and Klasebrun,M.
REFERENCE Direct Submission
TITLE Submitted (21-JUN-2000) Surgical Research, Children's Hospital, 320
JOURNAL Longwood Ave, Boston, MA 02115, USA
FEATURES
source
1..111934

```

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<3518. >111422
/gene="NRP2"
/gene="NRP2"
mrna
join(<3518. .3590,17760. .17937,36403. .36584,42688. .42918
43995. 44150,46123. .46292,48101. .48256,60728. .60872,
63412. .63761,65955. .66039,69934. .70050,73043. .73183,
83879. .84141,85679. .85775,86988. .87023,108610. .108660,
111118. >111422)

```

/gene="NRP2"
 /product="neuropilin-2a (12)"
 join(3518,.3590,17760,.17937,36403,.36584,42688,.42918
 43995,.44150,46123,.46292,48101,.48256,60728,.60872,
 63412,.63761,65955,.66099,69934,.70050,73043,.73183,
 83879,.84141,85679,.85775,86988,.87008,108610,.108660,
 11118,.111422)
 /gene="NRP2"
 /product="neuropilin-2a (17)"
 join(3518,.3590,17760,.17937,36403,.36584,42688,.42918,
 43995,.44150,46123,.46292,48101,.48256,60728,.60872,
 63412,.63761,65955,.66099,69934,.70050,73043,.73183,
 83879,.84141,85679,.85775,86988,.87023,108610,.108660,
 11118,.111422)
 /gene="NRP2"
 /note="receptor for VEGF165 and semaphorins class3"
 /codon_start=1
 /product="NRG1-2a (12)"
 /protein_id="AA61898.1"
 /db_xref="GI:11934949"
 /translation="MDMPEPLTWFLALYFSRHOVGRQDPDPCGRLNSKADGYITSPG
 YPODPSPHONCEMIVVAPENPQKIVLNFNPFIEHGDKYFIRIGDESADLLG
 KHCNGIAPETIISGSMLYIKFTSDYRQAGSLYEYLFKSGSDCKNFTSPGRT
 ESPPPEPRPNLDCFTTILAKPMELIIOPLIFDLHEHPLDVGSGDKYOMLIWIG
 IYHGYPLRYGCKTPKPSSELRSTGILSTFTPDMAVADGSAAYYLVHDEPLENFO
 CNVPLGMSGRIANEQISASSTYSDBRWTPQOORSLHGDNGTETPAUDSKYELVQDILR
 FLTM/LATQASLRETQNGYVYKSELESTNGEDMNVYHGNHNVFQANLADTE
 VLVN/LHAPLTLTRVIRPQTMHSGALRIELFCGVATDAPCSNNLGMISGLIADSOI
 SASSTOELVQSGSARLVSRSGCFWPIPAQOGEEMLVQDQPKTYKAYVIGQARG
 GDSITAYEARAFVRKPVSYSLNGKMEYDIOORPQOPLFGMNHVYTPPIRFDPI
 PAQVARYPERMSPAIGCMRLYLQCDMDTSKPYETTLGPTKYAKSETTTPYTEEAT
 ECGENCSFEDDKLDQPSGFNCNDFLEEECPGMMYDHAWMLTKTVAASSPNDRTPE
 DRNLTARLOSQOREQOARLISPVHLPSPVCMERQYATGGRVALQVYRASAQES
 KLVNIVIEDQCGEKGRILTPSYDMEOYIVFEGVYIGKRSGEIADIRISTVPLE
 NCMPEISAFANDEKFDVPEIHERREGYDEIDDEYEDVMSNASSATGSAPTDSEKMLY
 LDP/LITLITMSLGVLTAGTCAGLLCYTCYSYGLSSRCTLENNFELYDGLKHK
 VKNKHGCSSEA"
 join(43518,.3590,17760,.17937,36403,.36584,42688,.42918
 43995,.44150,46123,.46292,48101,.48256,60728,.60872,
 63412,.63761,65955,.66099,69934,.70050,73043,.73183,
 83879,.84141,85679,.85775,86988,.87023,96451,.96731)
 /gene="NRP2"
 /product="neuropilin-2b (5)"
 join(43518,.3590,17760,.17937,36403,.36584,42688,.42918
 43995,.44150,46123,.46292,48101,.48256,60728,.60872,
 63412,.63761,65955,.66099,69934,.70050,73043,.73183,
 83879,.84141,85679,.85775,86988,.87023,96451,.96731)
 /gene="NRP2"

```

63412. .63761.65955. .66099.69934. .70050.73043. .73183.
83879. .84141.85679. .85775.86988. .87008.96451. .96731)
/gene="NR2"
/product="neuropilin-2b(0)"
join(3518. .3590.17760. .17937.36403. .36584.42688. .42918
43995. .44150.46123. .46292.48101. .48256.60728. .60872.
63412. .63761.65955. .66099.69934. .70050.73043. .73183.
83879. .84141.85679. .85775.86988. .87008.96451. .96731)
/gene="NR2"
/note="receptor for VEGF165 and semaphorins class3"
/codon_start=1
/product="neuropilin-2b(5)"
/protein_id="A041900.1"
/db_xref="GI:11934951"
/translation="MDMFPLTWVFLALYFSRHOVRGQDPKPCGRLNSKDAGYITSPG
YPOYPHONCEWIVYAEPENOKIVLNFHFEIEXHDKYDFEIRDGSBADLLG
KHGNIAPPTIISGSMLYIKPTSDYARQAGSLYELFKTGSBDCSKYDPSNPTI
ESRPFREKYPHNLDCTTTLAKRMEIILQPLFLDHPDLQVGBEDCKYDLDIDNG
IIPHGLIKYCGCTKTPSELRSSTGSLTFPHDMDVAADGFSARYLHVOELENFO
VLNKHAFLPTFRVIRPOTMHSIALRELFGCVTAPKGMGLSGLAANDOI
SASTOELVMSPSAARLYVSRSGMPRIPOAGGEEMLDVDTGPTVYGVIIQARG
GDSITTYEARAPFRKRVYSISLNGKMEYI QDRTPQKPLFESNMHYDPIDRFPPI
PAQIVRYPERMSPAIGKMLEVLGCDWDSKTYTTLGPTVASETTTPIPREEAT
ECNGCSFDDKQOLPSEGNCFNDFLEPCGMVYHAKMLRTYMASSSPNDRTPD
BENRLASDOSOREGOYALRISPPNPLPRSPVCEYOYATGRGVALYVRBASDES
KLAWIREDOGGEWKGRILLPSEYMEYOIVFEGVIGKGSGLAIDLIRISVDLE
NOMEPRISAPAGENPKGCTILRGTEPVTDPVMOPIPAYMYVMAAGAVLVIVSVALA
LVLHFRYAAKTDHSITYKTSHTNGAPLAVEPTLTIKLQDGSHC"
join(3518. .3590.17760. .17937.36403. .36584.42688. .42918
43995. .44150.46123. .46292.48101. .48256.60728. .60872.
63412. .63761.65955. .66099.69934. .70050.73043. .73183.
83879. .84141.85679. .85775.86988. .87008.96451. .96731)
/gene="NR2"
/note="receptor for VEGF165 and semaphorins class3"
/codon_start=1
/product="neuropilin-2b(0)"
/protein_id="A041899.1"
/db_xref="GI:11934950"
/translation="MDMFPLTWVFLALYFSRHOVRGQDPKPCGRLNSKDAGYITSPG
YPOYPHONCEWIVYAEPENOKIVLNFHFEIEXHDKYDFEIRDGSBADLLG
KHGNIAPPTIISGSMLYIKPTSDYARQAGSLYELFKTGSBDCSKYDPSNPTI
ESRPFREKYPHNLDCTTTLAKRMEIILQPLFLDHPDLQVGBEDCKYDLDIDNG
IIPHGLIKYCGCTKTPSELRSSTGSLTFPHDMDVAADGFSARYLHVOELENFO
CNPVLGESGRIANEOISASSTYSIDGRMTPOGRLHGDNDGMPNLSKRYLVOLR
FLMULTRAIATQGLISRETONGYVYSYKLFVSTNGEDKNYVRHGNKHYFOANDNTE
VLNKLHAFLPTFRVIRPOTMHSIALRELFGCVTAPCSNMLGMSGLAANDOI
SASTOELVMSPSAARLYVSRSGMPRIPOAGGEEMLDVDTGPTVYGVIIQARG
GDSITTYEARAPFRKRVYSISLNGKMEYI QDRTPQKPLFESNMHYDPIDRFPPI
PAQIVRYPERMSPAIGKMLEVLGCDWDSKTYTTLGPTVASETTTPIPREEAT
ECNGCSFDDKQOLPSEGNCFNDFLEPCGMVYHAKMLRTYMASSSPNDRTPD
BENRLASDOSOREGOYALRISPPNPLPRSPVCEYOYATGRGVALYVRBASDES
KLAWIREDOGGEWKGRILLPSEYMEYOIVFEGVIGKGSGLAIDLIRISVDLE
NOMEPRISAPAGTLLPOTEPVTDPVMOPIPAYMYVMAAGAVLVIVSVALAVLHY
HRRRYAAKTDHSITYKTSHTNGAPLAVEPTLTIKLQDGSHC"

```


Db 77259 GCCCTCCCTCGGAACGAGAGCTGGGACCTCCCTCGGACCTCACCCTCATTTGGAA 77200

64 GGAGGTGAGGAGCCTCTGAGGAGGAGGCA 94

Db 77399 TGATGAGAGGTGTCTCCCAAGGCTGGGCA 77169

RESULT 12
AC159874/c
LOCUS
DEFINITION Bos taurus clone CH240-5418, *** SEQUENCING IN PROGRESS ***, 26
AC159874 233561 bp DNA linear HTG 01-JUL-2005
unordered pieces.
AC159874
GI:68267115
HTG, HTGS_PHASE1, HTGS_DRAFT, HTGS_ENRICHED.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Bovidae; Bovinae; Bos.
Munry, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D.,
Aryal-Becchi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Bernalto, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, K., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Claveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, J.,
Devila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Deigado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falle, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Freese, C. M., Gabisi, A., Garcia, R., Garcia, A., Garner, T., Garza, M.,
Georgescu, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guayra, W.,
Gunsberg, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kovis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenz, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindrasekaran, M., Mahmood, M., Malloy, K., Mangun, A.,
Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Muidana, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokolehen, O., Okunola, G., Olampong, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L. L.,
Puzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M. A., Reigh, R.,
Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C. D., Sosa, J. D.,
Snead, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J. D.,
Steinle, M., Strong, R., Sutton, A., Svatik, A., Tabors, Z., Uemami, K.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K.,
Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlecyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, S., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Zhao, S., Dunn, D., von
Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G., and Gibbs, R. A.

TITLE
JOURNAL
REFERENCE
2 (bases 1 to 233561)
Direct Submission
Unpublished

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Worley, K. C.
Direct Submission
Submitted (13-APR-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 233561)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 28, 2005 this sequence version replaced gi:62543289.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/atlas/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FDDS
Center clone name: CH240-5418
----- Summary Statistics
Assembly program: Atlas 3.0.
Consensus quality: 226744 bases at least Q40
Consensus quality: 228422 bases at least Q30
Consensus quality: 229869 bases at least Q20
Estimated insert size: 229962; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 26 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1
4475: contig of 4475 bp in length
4476
4794: gap of 319 bp
4795
10004: contig of 5210 bp in length
10005
10054: gap of 50 bp
10055
14989: contig of 4935 bp in length
14990
15039: gap of 50 bp
15040
23193: contig of 8153 bp in length
23193
23242: gap of 50 bp
23243
27656: contig of 4414 bp in length
27657
27707: gap of 50 bp
27707
27706: contig of 4639 bp in length
32346
32395: gap of 50 bp
32396
80065: contig of 47670 bp in length
80066
80258: gap of 193 bp
80259
95526: contig of 15268 bp in length
95527
95576: gap of 50 bp
95577
11151: contig of 15575 bp in length
11152
111201: gap of 50 bp
111202
127875: contig of 16674 bp in length
127876
127925: gap of 50 bp
127926
130975: contig of 3050 bp in length
130976
130976
136494: contig of 5430 bp in length

* 136495 136594: gap of unknown length
* 136595 140341: contig of 3747 bp in length
* 140342 140391: gap of 50 bp
* 140392 141698: contig of 1307 bp in length
* 141699 141748: gap of 50 bp
* 141749 151343: contig of 9595 bp in length
* 151344 151393: gap of 50 bp
* 151394 154790: contig of 3397 bp in length
* 154791 154840: gap of 50 bp
* 154841 166899: contig of 12059 bp in length
* 166900 172623: contig of 5674 bp in length
* 172624 172673: gap of 50 bp
* 172674 185758: contig of 13085 bp in length
* 185759 185808: gap of 50 bp
* 185809 210143: contig of 24335 bp in length
* 210144 210193: gap of 50 bp
* 210194 226003: contig of 15810 bp in length
* 226004 226103: gap of unknown length
* 226104 227673: contig of 1570 bp in length
* 227674 227773: gap of unknown length
* 227774 229134: contig of 1361 bp in length
* 229135 229234: gap of unknown length
* 229235 230327: contig of 1093 bp in length
* 230328 230427: gap of unknown length
* 230428 231641: contig of 1214 bp in length
* 231642 231741: gap of unknown length
* 231742 233561: contig of 1820 bp in length.

FEATURES

SOURCE

1..233561
Location/Qualifiers
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-5418"
4476..4794
/estimated_length=319
10005..10054
/estimated_length=50
14990..15039
/estimated_length=50
23193..23242
/estimated_length=50
27657..27706
/estimated_length=50
32346..32395
/estimated_length=50
80066..80258
/estimated_length=193
95527..95576
/estimated_length=50
11152..11201
/estimated_length=50
127876..127925
/estimated_length=50
130976..131064
/estimated_length=89
136495..136594
/estimated_length=unknown
140342..140391
/estimated_length=50
141699..141748
/estimated_length=50
151344..151393

Query Match 7.1%; Score 39.8; DB 14; Length 233561;
Best Local Similarity 60.7%; Pred. No. 33;
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 274 GCCTTGACGGGTGGGCTTCACGACAGGGTCCCATCCACGATCCAGCCCAAGTGGC 333
DB 125658 GCTTGCCCGAGGATTAAGCTGTGCTGCTGAGGCTGAGCCCCACGTCACGACATGCCCGC 125599
QY 334 TGACAGAGGTCCCTCGCAGTCAATGAACCAAGGAGGCTTGGAAC 380

Db 125598 TCTGCAAGTCCCTCAGGCTCACTGTCAGAGGTGGGGGTGGGAC 125552

RESULT 13
AC138582/C
LOCUS AC138582 119613 bp DNA linear PRI 04-AUG-2005
DEFINITION Pan troglodytes clone rp43-21p1, complete sequence.
AC138582
ACCESSION
VERSION AC138582.22 GI:71067221
KEYWORDS
SOURCE HTG.
ORGANISM Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pan.

REFERENCE
AUTHORS Lau,C.C.Y. and Roe,B.A.
TITLE Pan troglodytes BAC Clone rp43-21p1
JOURNAL Unpublished
REFERENCE
AUTHORS Lau,C.C.Y. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS Lau,C.C.Y. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2005) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

On Jul 23, 2005 this sequence version replaced gi:67514672.

FEATURES
source
1..119613
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="rp43-21p1"
/clone_lib="RPCI - 43 Male Chimpanzee BAC Library"

ORIGIN
Query Match 7.0%; Score 39.6; DB 8; Length 119613;
Best Local Similarity 46.9%; Pred. No. 40;
Matches 123; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 3 AGCCCTGCTCAGAAACAGAAAGGACGAGACACTCAGCGTGACTACCCCATGTGGC 62
DB 85272 AGCTGAGAGCAACCTGGGAGGACCTGCTGGCTGTGATGGAGTACAGAGG 85213
QY 63 TGAAGTGAAGAGCCTCTGAGGACGAGGACGAGGACCGTCAAGTGGTGAACGAGG 122
DB 85212 CCTAGTGAAGGAGGAGCATGAGGACGAGGCTTCAGAGAGACGGGGTGGCATTTGGGAG 85153
QY 123 GGTCTTGCCATGTGGGACAGGAGGCTGCATACAGCTTACTCACTGATGACATCGACTCCT 182
DB 85152 GTCATGGGCAAGTGGGAGGTCAAGCTGTGAGCACCTTCCCAAGGCCAATGACCAAAAT 85093
QY 183 GGTGCACGCTCTGGAATCTGGAAGTGAAGCAATGTTCCATTAAGGAAGTGTGGC 242
DB 85092 GAGACAACTTTAAGACTGACAGACATGTGGGGGCTGTGGAGCCAGGAAGCTAATGTTC 85033

DB	85032	ACACAACACCCCGAGCTCTGC	85011
RESULT 14	AC000029/c	142979 bp	DNA linear PRI 09-APR-2003
LOCUS	AC000029/c	142979 bp	DNA linear PRI 09-APR-2003
DEFINITION	Complete sequence.	22q11.2 BAC clone 865e9 in GNA2-BCR region	
ACCESSION	AC000029		
VERSION	AC000029.17		
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	Kim, U.-J. and Simon, M.		
JOURNAL	1 (bases 1 to 142979)		
REFERENCE	Unpublished		
AUTHORS	Wang, Z. and Roe, B.A.		
JOURNAL	2 (bases 1 to 142979)		
REFERENCE	Unpublished		
AUTHORS	3 (bases 1 to 142979)		
JOURNAL	Roe, B.A.		
REFERENCE	Submitted (21-OCT-1996) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
AUTHORS	4 (bases 1 to 142979)		
JOURNAL	Roe, B.A.		
REFERENCE	Submitted (13-SEP-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
AUTHORS	5 (bases 1 to 142979)		
JOURNAL	Roe, B.A.		
REFERENCE	Submitted (24-SEP-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
AUTHORS	6 (bases 1 to 142979)		
JOURNAL	Roe, B.A.		
REFERENCE	Submitted (01-JUN-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
AUTHORS	7 (bases 1 to 142979)		
JOURNAL	Roe, B.A.		
REFERENCE	Submitted (09-APR-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
AUTHORS	On Sep 24, 1999 this sequence version replaced gi:5870976.		
JOURNAL	Because these overlapping clones came from different libraries.		
COMMENT	Location/Qualifiers		
FEATURES	1..142979		
SOURCE	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="22"		
	/map="22q11.2"		
	/clone="865e9"		
ORIGIN			
	Query Match	7.0%;	Score 39.6; DB 8; Length 142979;
	Best Local Similarity	46.9%;	Pred. No. 39;
	Matches 123; Conservative	0;	Mismatches 139; Indels 0; Gaps 0;
3	AGCCCTGCTCAGAAACAGAGACGACGACATCTCAGCGTGATCTCACCCTCATGTGCG	62	

Db	133180	AGCTCAGACGAACTGCGGGAGGGCCACCCTGCTG6GCTGTGTA6TGGATCAGAGAG	133121
Qy	63	TGAGAGTGAAGGAGCCTCTCTGAGCGAGGGCCAGGGCAGCCCTGACAGTGGGTGAACGGCAG	122
Db	133120	CCTAGGTGAGGGAGGAGTGTGAGCCAGGGCCTTCAAGAGAGTGGGGTGGGATTTGGGGAG	133061
Qy	123	GGTCTTGGCATAGTGTGGGACAGGGGGCTGCAATACAGTTACTAGTGAACAATGAGTCCCT	182
Db	133060	GTTCATGGCGAGGTGGGACAGGTCAAGGCTGTGAGCACTTTCCTCCAGACCATGACCAAAAT	133001
Qy	183	GGTGCACAGCCTCTGGAAGTCTGGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTGTGC	242
Db	133000	GAGACAACTTTTGGGACTAGCAGAGAGATGTGGGGGCTCTGGGAGCCAGGAAGCTAAATGTG	132941
Qy	243	CGCCATGCCCCCAACGTTGC	264
Db	132940	ACACAAACACCCCAAGCTCTGC	132919
RESULT 15			
LOCUS	AC148588		
DEFINITION	AC148588	190018 bp	DNA
ACCESSION	AC148588		linear
VERSION	AC148588.26	GI:71067215	HTG 23-JUL-2005
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.		
SOURCE	Pan troglodytes (chimpanzee)		WORKING DRAFT SEQUENCE, 3 ordered
ORGANISM	Pan troglodytes		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Pan.		
REFERENCE	1	(bases 1 to 190018)	
AUTHORS	Lau,C., Fujiyama,A. and Roe,B.A.		
TITLE	Pan troglodytes BAC Clone pcb-92Fz3		
REFERENCE	Unpublished		
AUTHORS	2	(bases 1 to 190018)	
TITLE	Lau,C., Fujiyama,A. and Roe,B.A.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (18-MAR-2004)		Department Of Chemistry And Biochemistry,
TITLE	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,		
REFERENCE	OK 73019, USA		
AUTHORS	3	(bases 1 to 190018)	
TITLE	Lau,C., Fujiyama,A. and Roe,B.A.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (23-JUL-2005)		Department Of Chemistry And Biochemistry,
TITLE	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,		
REFERENCE	OK 73019, USA		
COMMENT	On Jul 23, 2005 this sequence version replaced gi:66473101.		
	----- Genome Center		
	Center: Department Of Chemistry And Biochemistry		
	The University Of Oklahoma		
	Center code:UOKNOR		

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 3 contigs. Gaps between the contigs		
	* are represented as runs of N. The order of the pieces		
	* is believed to be correct as given, however the sizes		
	* of the gaps between them are based on estimates that have		
	* provided by the submitter.		
	* This sequence will be replaced		
	* by the finished sequence as soon as it is available and		
	* the accession number will be preserved.		
	1 137286: contig of 137286 bp in length		
	* 137287 137286: gap of unknown length		
	* 137387 145329: contig of 7943 bp in length		
	* 145330 145429: gap of unknown length		
	* 145430 190018: contig of 44589 bp in length.		
	Location/Qualifiers		
	1. 190018		
FEATURES	/organism="Pan troglodytes"		
source	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		

gap /clone="ptb-94f23"
gap /clone_lib="PTB Male Chimpanzee BAC Library"
gap /estimated_length=unknown
gap /estimated_length=unknown
gap /estimated_length=unknown

ORIGIN

Query Match 7.0%; Score 39.6; DB 14; Length 190018;

Best Local Similarity 46.9%; Pred. No. 38; Mismatches 139; Indels 0; Gaps 0;

Matches 123; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 3 AGCCCTGCCTCAGAAAAACGAAAGACGACACTCAGCGTGACTCAACCCCATGTGGC 62
DB 165064 AGCTCAGCAGCAACCTGCGGAGGCGCACCCCTGTGGCTGTGAGTGGGATCACAAGAG 165123
QY 63 TGGAGGTGAGGAGCCTCTCTGAGGCAAGGCCAGGGCAGCCGTCAGGTGGGTACCGCAG 122
DB 165124 CCTAGGTCAAGGAGGACTGAGGCCAGGGCCTCGAAGAGAGCGGGGTGGCAATTGGGGAG 165183
QY 123 GGTCTTGCCATGATGGGACAGGGGCTGCATACAGCTTACTCAGTGACAAATCGAGTCCCT 182
DB 165184 GTCATGGGAGAGGTGGGACAGGTACAGGCTGTGAGCACTTCCCGAGGCCAATGACCAAAAT 165243
QY 183 GGTGCCAGCCTCTGAAAGTCTGAAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTGGC 242
DB 165244 GAGACAACCTTTAGACTGACAGACATGTGGGGGCTCTGGAGCCAGAAAGCTTAAGTGTG 165303
QY 243 CGGCCATGCCCCCAACGTTGC 264
DB 165304 ACACAACACCCCAAGGCTCTGC 165325

Search completed: March 25, 2006, 17:51:44
Job time : 3011.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 15:57:57 ; Search time 472 Seconds
(without alignments)
7963.743 Million cell updates/sec

Title: SEQ1-69T
Perfect score: 564
Sequence: 1 gcagccctgcctcagaaac.....agtcgctcttcgacttc 564

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*
14: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	563	99.8	564	13	ADT97901
2	551.4	97.8	29340	13	ABD33482 Human fib
3	551.4	97.8	57501	14	AD213482
4	38.2	6.8	2161	8	ABX34807 Human can
5	38.2	6.8	2777	11	ADM01893 Human mdd
6	37.2	6.6	712	6	ABK34393 Human CDN
7	37.2	6.6	1957	4	ABK42424 Genomic s
8	37.2	6.6	1957	4	AA107453 Human rep
9	37.2	6.6	1957	9	ADB60580 Connectiv
10	36.2	6.4	138203	14	AD251725 FR-008 po
11	36	6.4	546	13	ADQ56739
12	35.8	6.3	704	2	AAQ94044 Human ALD
13	35.8	6.3	1263	6	ABQ90455 M. capsul
14	35.6	6.3	568	5	ABV51994 Human pro
15	35.4	6.3	1785	6	AAI51911 Reference
16	35.2	6.2	954	14	ACT170503 M. xanthu
17	35.2	6.2	4050	14	ACL64141 M. xanthu
18	35	6.2	550	9	ACH25075 Human adu
19	35	6.2	1338	6	ABL58452 Human pel

20	35	6.2	1338	12	ADP48674	Adp48674 Human pel
21	35	6.2	1797	4	ABK42425	ABK42425 Genomic s
22	35	6.2	1797	4	AA107455	AA107455 Human rep
23	35	6.2	1797	9	ADB60581	ADB60581 Connectiv
24	34.8	6.2	792	13	ABD62531	ABD62531 Cotton cd
25	34.8	6.2	76180	13	ABD33385	ABD33385 Human can
26	34.6	6.1	40742	4	AAK68089	AAK68089 Human imm
27	34.6	6.1	40742	4	AAK79886	AAK79886 Human imm
28	34.6	6.1	225980	6	AA138337	AA138337 Complemen
29	34.6	6.1	220895	6	ABK84798	ABK84798 Human CDN
30	34.6	6.1	220895	13	ADK52737	ADK52737 Drug ther
31	34.4	6.1	5051	13	ADQ39151	ADQ39151 Human SNP
32	34.4	6.1	5232	13	ADQ39152	ADQ39152 Human SNP
33	34.4	6.1	92638	6	ABO88096	ABO88096 Human ost
34	34.2	6.1	756	8	AAK85016	AAK85016 Human sec
35	34.2	6.1	756	8	ACD18942	ACD18942 Novel hum
36	34.2	6.1	756	12	ADG78333	ADG78333 Human sec
37	34.2	6.1	756	12	ADN60624	ADN60624 Human sec
38	34.2	6.1	349980	6	ABQ81844	ABQ81844 Bifidobac
39	34	6.0	593	12	ACH71544	ACH71544 Human gen
40	34	6.0	1028	10	ADD84538	ADD84538 121P1 v
41	34	6.0	1028	10	ADD84540	ADD84540 121P1 v
42	34	6.0	1028	12	ADM83831	ADM83831 Human can
43	34	6.0	1028	12	ADM83794	ADM83794 Human can
44	34	6.0	1028	12	ADM83796	ADM83796 Human can
45	34	6.0	1988	13	ADS56480	ADS56480 Bacterial

ALIGNMENTS

RESULT 1	ADT97901	ADT97901 standard; DNA, 564 BP.
ID	ADT97901	
XX	ADT97901;	
AC	27-JAN-2005	(first entry)
DT	Human fibroblast growth factor (FGF)-3 promoter 5' proximal region.	
DB	Human, ds; fibroblast growth factor-3; FGF-3; promoter; SNP;	
KW	single nucleotide polymorphism; cancer; oesophageal cancer; 5'UTR;	
KW	breast cancer; ovarian cancer; prostate cancer; head and neck cancer;	
KW	oesophageal squamous cell carcinoma.	
XX	Homo sapiens.	
OS		
XX		
PH	Key	Location/Qualifiers
FT	variation	replace(69,Y)
FT		/*tag= a
FT		/standard_name= "Single nucleotide polymorphism"
FT		/note= "This SNP is specifically claimed in claim 6"
XX	US2004219582-A1.	
XX	04-NOV-2004.	
XX	11-MAR-2004; 2004US-00798652.	
PF	17-MAR-2003; 2003US-0455689P.	
XX	(GUOY/) GUO Y.	
XX		
XX	Guo Y;	
XX	WPI; 2004-794435/78.	
DR		
XX		
FT	Novel isolated nucleic acid molecule having single nucleotide	
FT	polymorphism in upstream untranslated region of fibroblast growth factor-	
PT	3 gene, useful for assessing related susceptibility of mammal to cancer.	
XX	Claim 1; SEQ ID NO 1; 15pp; English.	

XX The invention relates to an isolated nucleic acid molecule comprising the
CC upstream untranslated region (UTR) of fibroblast growth factor (FGF)-3
CC gene sequence (AD97901) with a single nucleotide polymorphism (SNP) at
CC position 69. Also included are an isolated nucleic acid molecule
CC comprising a sequence complementary to AD97901, a vector comprising
CC AD97901 operably linked to a reporter gene, a host cell comprising the
CC vector, detecting a SNP in the FGF-3 gene in a mammal (involves isolating
CC a nucleic acid sample from the mammal, and determining whether a cytosine
CC or thymine is present at position 69 of AD97901), a kit for performing
CC the method (comprising a first oligonucleotide probe which anneals
CC specifically with the target portion of the mammal's genome, where the
CC first probe comprises a first fluorescent label and a first fluorescence
CC quencher attached to its separate nucleotide residues and the target
CC portion includes the nucleotide residue located at position 69 of
CC AD97901, and a pair of primers for amplifying a reference portion of the
CC FGF-3 gene, where the reference portion includes the nucleotide residue
CC located at position 69 of AD97901) and a microarray having at least one
CC oligonucleotide probe that can anneal with a target portion of a mammal's
CC genome, where the target portion includes the nucleotide residue located
CC at position 69 of AD97901. The method is useful for detecting SNP in FGF
CC -3 gene in a mammal, preferably a human, and is also useful for assessing
CC the relative susceptibility of a mammal to cancer (especially oesophageal
CC cancer), which shows an association with the presence of the C-allele.
CC The cancer is chosen from oesophageal, breast, ovarian, prostate, head
CC and neck cancer. The oesophageal cancer is oesophageal squamous cell
CC carcinoma. The present sequence is the upstream untranslated region (UTR)
CC of fibroblast growth factor (FGF)-3 gene sequence.

XX Sequence 564 BP; 117 A; 160 C; 185 G; 101 T; 0 U; 1 Other;

XX Query Match 99.8%; Score 563; DB 13; Length 564;

XX Best Local Similarity 99.8%; Pred. No. 2,3e-142;

XX Matches 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCAGCCCTGCTCAGAAAACAGAGAGCGACGACACTCAGCGTACATCCCATGTG 60
DB 1 GCAGCCCTGCTCAGAAAACAGAGAGCGACGACACTCAGCGTACATCCCATGTG 60
QY 61 GGTGAGGTGAGAGAGGCTCTGAGGCGAGGCGAGGCGCGGTGAGGGGTGACGCA 120
DB 61 GCTGAGAGGAGAGAGGCTCTGAGGCGAGGCGAGGCGCGGTGAGGGGTGACGCA 120
QY 121 GGGGCTTGGCATGTGGGCAACAGGGGCTGATACAGCTTACTCAGTGCATCGAGTCC 180
DB 121 GGGGCTTGGCATGTGGGCAACAGGGGCTGATACAGCTTACTCAGTGCATCGAGTCC 180
QY 181 CTGGTGCCAGGCTCTGGAAGTGTGAAGTGAACAATGTTTCCATTAAAGAAAGTGTG 240
DB 181 CTGGTGCCAGGCTCTGGAAGTGTGAAGTGAACAATGTTTCCATTAAAGAAAGTGTG 240
QY 241 GCGGCGCATGCCCCCAAGTGGCACTCAGCTGCTTGGCAGGGGTGGGCTTCCAGTC 300
DB 241 GCGGCGCATGCCCCCAAGTGGCACTCAGCTGCTTGGCAGGGGTGGGCTTCCAGTC 300
QY 241 GCGGCGCATGCCCCCAAGTGGCACTCAGCTGCTTGGCAGGGGTGGGCTTCCAGTC 300
DB 241 GCGGCGCATGCCCCCAAGTGGCACTCAGCTGCTTGGCAGGGGTGGGCTTCCAGTC 300
QY 301 ACAGGGTCCCATCAAGTACAGTCCAGGCTGAGAGAGGTCCTCGCAGTCATGAAA 360
DB 301 ACAGGGTCCCATCAAGTACAGTCCAGGCTGAGAGAGGTCCTCGCAGTCATGAAA 360
QY 361 CCAAGGAGGCTTGGGAAAACAATCTGAAGGGCATGCGTTTGAATTAAGTGAAGAGGTG 420
DB 361 CCAAGGAGGCTTGGGAAAACAATCTGAAGGGCATGCGTTTGAATTAAGTGAAGAGGTG 420
QY 421 GCGTGGGCTGGGCAAGGCCACAGGCTGTGATCAAGAGCCAGGAGGAGAGTGTGCTCC 480
DB 421 GCGTGGGCTGGGCAAGGCCACAGGCTGTGATCAAGAGCCAGGAGGAGAGTGTGCTCC 480
QY 481 AGCAGTGCAGCGGCTCTGGATGAGTCACTCTGCGCACTGAGAAACAGCCTTGAAG 540
DB 481 AGCAGTGCAGCGGCTCTGGATGAGTCACTCTGCGCACTGAGAAACAGCCTTGAAG 540
QY 541 AGGAGTGGCGCTCTTTCGAGCTTC 564
DB 541 AGGAGTGGCGCTCTTTCGAGCTTC 564

DB 541 AGGAGTGGCGCTCTTTCGAGCTTC 564

RESULT 2

ABD33482

ABD33482 standard; DNA; 29340 BP.

18-NOV-2004 (first entry)

Human cancer-associated (CA) gene HD07-093.

Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;

de; cancer; cytostatic.

Homo sapiens.

WO2004058146-A2.

15-JUL-2004.

15-DEC-2003; 2003WO-US040081.

17-DEC-2002; 2002US-00322281.

(SAGR-) SAGRES DISCOVERY INC.

Morris DW, Malandro MS;

WPI; 2004-499109/47.

Novel human cancer associated protein encoded within open reading frame
of cancer associated gene, useful as targets for diagnosing cancer.

Clalm 16; SEQ ID NO 642; 182pp; English.

The invention relates to cancer-associated proteins (CAP) and the cancer-
associated (CA) nucleic acids encoding them. The invention also relates
to a method for treating cancers involving administering to a patient an
inhibitor of CAP, and a method of screening for anticancer activity in a
potential drug involving providing a cell that expresses a CA gene,
contacting a tissue sample derived from a cancer cell with an anticancer
drug candidate and monitoring the effect of the anticancer drug candidate
on expression of the CA gene. The CAP proteins are useful for detecting
cancer associated with expression of a CAP protein in a test cell sample
and for screening for a bioactive agent capable of modulating the
activity of a CAP protein. The CA nucleic acids are useful for diagnosing
cancer, involving determining the expression of a CA nucleic acid in a
tissue. This sequence represents a human CA gene of the invention. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences

Sequence 29340 BP; 6351 A; 8588 C; 8017 G; 6384 T; 0 U; 0 Other;

Query Match 97.8%; Score 551.4; DB 13; Length 29340;

Best Local Similarity 99.6%; Pred. No. 9.3e-139;

Matches 563; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GCAGCCCTGCTCAGAAAACAGAGAGCGACGACACTCAGCGTACATCCCATGTG 60
DB 1 GCAGCCCTGCTCAGAAAACAGAGAGCGACGACACTCAGCGTACATCCCATGTG 60
QY 61 GCTGAGGTGAGAGAGGCTCTGAGGCGAGGCGAGGCGCGGTGAGGGGTGACGCA 120
DB 61 GCTGAGGTGAGAGAGGCTCTGAGGCGAGGCGAGGCGCGGTGAGGGGTGACGCA 120
QY 121 GGGGCTTGGCATGTGGGCAACAGGGGCTGATACAGCTTACTCAGTGCATCGAGTCC 180
DB 121 GGGGCTTGGCATGTGGGCAACAGGGGCTGATACAGCTTACTCAGTGCATCGAGTCC 180
QY 181 CTGGTGCCAGGCTCTGGAAGTGTGAAGTGAACAATGTTTCCATTAAAGAAAGTGTG 240
DB 181 CTGGTGCCAGGCTCTGGAAGTGTGAAGTGAACAATGTTTCCATTAAAGAAAGTGTG 240

PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232860P.
PR 08-SEP-2000; 2000US-0232861P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234976P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251866P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-565190/63.

XX Nucleic acid encoding novel connective tissue associated polypeptides,
XX used in diagnosing, preventing, treating or ameliorating a disorder such
XX as cancer or rheumatoid arthritis.

XX Disclosure: SEQ ID NO 1311; 673pp; English.

PS The present invention relates to the isolation of novel human connective
XX tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
XX (cDNA and genomic) sequences encoding them. The sequences of the
XX invention are useful in the diagnosis, treatment, prevention and/or
XX prognosis of diseases associated with connective tissue(s), including
XX cancer. The polynucleotide sequences of the invention are also useful in
XX gene therapy. ABR42102-ABR43116 represent genomic sequences encoding the
XX novel human connective tissue related polypeptides. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1957 BP; 515 A; 580 C; 441 G; 421 T; 0 U; 0 Other;

XX Query Match 6 6%; Score 37.2; DB 4; Length 1957;

XX Best Local Similarity 56.6%; Pred. No. 8.8; 53; Indels 0; Gaps 0;

XX Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 251 CCCCCAAGTTGACACTGCTTGGACAGGTTGGGCTTCCAGTCACAGGATCCC 310

Db 165 CCGCCATGCTGCTGCTCACTTCTGCTCCCTGAGGAAAGGCGCGGCTTCTGCTCC 224

QY 311 ATCAAGTACCAAGCCAGGTGCTGACAGAGTCTCTGCACTATGAAACCAAGGAGG 370

Db 225 AGGAATCTCCACACAGAGATCTGACAGAGGAGCTGCGCTGCACTAGTCCAGAAACA 284

QY 371 CT 372

Db 285 CT 286

RESULT 8
 ID AL07453
 AC AL07453;
 DT 21-NOV-2001 (first entry)
 DE Human reproductive system related antigen DNA SEQ ID NO: 10141.
 KW Human; reproductive system related antigen; reproductive system disorder;
 cancer; gene therapy; db.
 XX Homo sapiens.
 OS
 XX
 XX WO200155320-A2.
 PD 02-AUG-2001.
 XX
 PE 17-JAN-2001; 2001WO-US001339.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-021680P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226881P.
 PR 22-AUG-2000; 2000US-0226886P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227183P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229309P.
 PR 05-SEP-2000; 2000US-0229517P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0233080P.

PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246529P.
 PR 08-NOV-2000; 2000US-0246530P.
 PR 08-NOV-2000; 2000US-0246531P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.

[illegible]

FT	/tag= c
FT	/product= "regulator protein, FscR1, SEQ ID 10"
FT	complement(4377. .7205)
FT	/tag= d
FT	/product= "Regulator protein, FscRII, SEQ ID 11"
FT	complement(7210. .10320)
FT	/tag= e
FT	/product= "Regulator protein, FscRIII, SEQ ID 12"
FT	complement(10289. .13315)
FT	/tag= f
FT	/product= "Regulator protein, FscRIV, SEQ ID 13"
FT	13522. .14898
FT	/tag= g
FT	/product= "Glycosyltransferase, FscMI, SEQ ID 17"
FT	14953. .16011
FT	/tag= h
FT	/product= "GDP-ketosugar aminotransferase, FscMII, SEQ ID
FT	18"
FT	16106. .17287
FT	/tag= i
FT	/product= "Cytochrome P-450 monooxygenase, FscP, SEQ ID
FT	14"
FT	17334. .17528
FT	/tag= j
FT	/product= "Ferredoxin protein, FscFE, SEQ ID 15"
FT	17556. .18413
FT	/tag= k
FT	/product= "Phioesterase, FscTE, SEQ ID 16"
FT	18610. .20781
FT	/tag= l
FT	/product= "ADC synthase, PabAB, SEQ ID 21"
FT	20927. .26158
FT	/tag= m
FT	/product= "PKS FscA, SEQ ID 2"
FT	26333. .27340
FT	/tag= n
FT	/product= "ABC transporter, FscTI, SEQ ID 8"
FT	27561. .28280
FT	/tag= o
FT	/product= "ABC transporter, FscTII, SEQ ID 9"
FT	complement(28983. .60860)
FT	/tag= p
FT	/product= "PKS FscC, SEQ ID 3"
FT	complement(60962. .77587)
FT	/tag= q
FT	/product= "PKS FscB, SEQ ID 4"
FT	complement(77983. .84132)
FT	/tag= r
FT	/product= "PKS FscF, SEQ ID 5"
FT	complement(84170. .107485)
FT	/tag= s
FT	/product= "PKS FscE, SEQ ID 6"
FT	complement(107496. .136148)
FT	/tag= t
FT	/product= "PKS FscD, SEQ ID 7"
FT	complement(136558. .137766)
FT	/tag= u
FT	/product= "GDP-mannose-4,6-dehydratase, FscMIII, SEQ ID
FT	19"

DR WP1 : 2005-332033/33.
DR P-SDB: ADZ51726, ADZ51727, ADZ51728, ADZ51729, ADZ51730, ADZ51731,
DR ADZ51732, ADZ51733, ADZ51734, ADZ51735, ADZ51736, ADZ51737, ADZ51738,
DR ADZ51739, ADZ51740, ADZ51741, ADZ51742, ADZ51743, ADZ51744, ADZ51745,
XX ADZ51746.

XX
XX
PT New gene cluster for the biosynthesis of FR-008 polyketides, useful for
PT increasing the productivity of FR-008 polyketides or to produce new FR-
PT 008 polyketide variants.

PS Claim 1; SEQ ID NO 1; 22pp; English.

XX
CC The present invention relates to a gene cluster (ADZ51725) sp. for the
CC biosynthesis of FR-008 polyketides derived from Streptomyces sp. FR-008.
CC FR-008 polyketide, a heptaene macrolide having aglycone containing 4-
CC aminocyclohexanone, has antifungal activity and also high toxicity against
CC mosquito larvae, and so is highly useful in agricultural and medical
CC fields. The gene cluster comprises the following genes: fscA, fscC, fscB,
CC fscE, fscD and fscI, which code for modular polyketide synthase (PKS);
CC fscF, fscG and fscH, which code for modular polyketide synthase (PKS);
CC fscTII and fscTIII, which code for ABC transporter proteins; fscR1, fscRII,
CC fscRIII and fscIV genes which code for regulator proteins; fscP which
CC codes for cytochrome P450 monooxygenase; fscQf which codes for ferredoxin
CC protein; fscW which codes for thioesterase; fscM1 which codes for
CC glycoyltransferase; fscMII which codes for GDP-ketoylur
CC amino transferase; fscMIII which codes for GDP-mannose-4,6-dehydratase;
CC fscO which codes for PAD-dependent monooxygenase; pabAB which codes for 4
CC -amino-4-deoxychorismate (ADC) synthase and pabC which codes for ADC
CC lyase. The genes of the gene cluster of the invention can be used to
CC develop recombinant microorganisms capable of producing FR-008
CC polyketides, also to increase the productivity of the existing FR-008
CC polyketides or to produce new FR-008 polyketide variants, by its
CC modification.

XX Sequence 138203 BP; 18153 A; 48202 C; 55342 G; 16506 T; 0 U; 0 Other;

Query Match	6.4%	Score 36.2;	DB 14;	Length 138203;
Best Local Similarity	55.0%	Pred. No. 53;		
Matches 71; Conservative	0;	Mismatches 58;	Indels 0;	Gaps 0;

[illegible]

RESULT 11
ADQ56739
ID ADQ56739 standard; DNA; 546 BP.

DT 21-OCT-2004 (first entry)

DE Novel canine microarray-related DNA sequence SegID8041.

KM canine microarray; drug screening; toxicity assay;
KM environmental pollutant; cellular response; gene expression profile;
KM toxic response; liver necrosis; fatty liver disease;
protein adduct formation; hepatitis; dog; ds.

OS Canis familiaris.

PN WO2004063324-A2.

PD 29-JUL-2004.

XX

Job time : 473 secs

AC AAS19511;

XX 26-MAR-2002 (first entry)

XX Reference sequence for human MPL gene exons 9-10.

XX Human; single nucleotide polymorphism; SNP; MPL; chromosome 1p34;
 KW myeloproliferative leukemia virus oncogene; haplotyping; genotyping;
 KW congenital amegakaryocytic thrombocytopaenia; CAMT; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH 614..773
 FT /*tag= a
 FT exon

FT /number= 9
 FT 774..1027
 FT /*tag= b
 FT Intron

FT /number= 9
 FT 1028..1124
 FT exon
 FT /*tag= c
 FT /number= 10

XX W0200179232-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US012301.

XX 14-APR-2000; 2000US-0197839P.

XX (GENA-) GENAISSANCE PHARM INC.

XX Chew A, Choi JY, Koshy B, Stephens JC;

XX WPI; 2002-055251/07.

XX P-PSDB; AAU09495.

XX Nucleotide polymorphisms in the human myeloproliferative leukemia virus
 PT oncogene (MPL) gene, useful for studying the function of and expressing
 PT MPL protein for use in screening drugs for treating diseases related to
 PT MPL activity.

PS Claim 27; Fig 3; 85pp; English.

XX The present invention relates to novel single nucleotide polymorphisms
 CC (SNPs) in the human myeloproliferative leukemia virus oncogene (MPL)
 CC gene located on chromosome 1p34, and methods for haplotyping and/or
 CC genotyping the MPL gene. The methods of the invention make use of allele-
 CC specific oligonucleotides (ASOs) as probes and primers and/or primer-
 CC extension oligonucleotides for detecting MPL gene polymorphisms. The
 CC polynucleotides and screened compounds are useful for the treatment of
 CC diseases associated with MPL activity, such as congenital amegakaryocytic
 CC thrombocytopaenia (CAMT). The present sequence represents a reference
 CC sequence for human MPL gene exons 9-10. Note: This sequence encodes for
 CC only part of the MPL protein shown in AAU09495

XX Sequence 1785 BP; 317 A; 562 C; 500 G; 403 T; 0 U; 3 Other;

Query Match 6.3%; Score 35.4; DB 6; Length 1785;
 Best Local Similarity 61.3%; Pred. No. 26;
 Matches 57; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

OY 53 CCCATGTGCTGTAGAGTGAAGAGCTCTCTTGAAGCAGGCGGACAGCCCTCAAGTGGG 112
 DB 913 CGCAGGTGCTCCGACGTCCAGGCGGCGCAGAGGCGGCGGCGCAGAGTGAAGGCTGCTGA 972

OY 113 TGACGGCAGAGGGGTCTTGCCTGATGTGGCAGG 145
 DB 973 TGAGGCGGGGCTCCGGCCCGGCTGGGCTCGAAG 1005

Search completed: March 25, 2006, 16:28:08

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2006, 16:12:32 ; Search time 3244 Seconds
(without alignments)
8134.376 Million cell updates/sec

Title: SEQ1-69T
Perfect score: 564
Sequence: 1 gcagccctgcctcagaaac.....agtgccttcgcgacttc 564

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_ests1:
2: gb_ests2:
3: gb_ests3:
4: gb_hnc:
5: gb_ests4:
6: gb_ests5:
7: gb_ests6:
8: gb_ests7:
9: gb_gss1:
10: gb_gss2:
11: gb_gss3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128.4	22.8	603	AG133394	AG133394 Pan trogl
2	41	7.3	312	A1680295	A1680295 tw62a11.x
3	40.4	7.2	925	CNS0091P	AL053013 Drosophi1
4	40	7.1	925	CNS0091P	AL053013 Drosophi1
5	39.6	7.0	915	B0948345	B0948345 AGENCOURT
6	39.6	7.0	1101	CNS0178Y	AL108460 Drosophi1
7	38.8	6.9	321	A1289637	AL1289637 qw29d03.x
8	38.2	6.8	637	CL542202	CL542202 OB_Ba006
9	38.2	6.8	867	BQ931784	BQ931784 AGENCOURT
10	38	6.7	393	A1970105	A1970105 wq89c02.x
11	37.8	6.7	481	A1818367	A1818367 wks5h08.x
12	37.8	6.7	523	A1818367	A1818367 wks5h08.x
13	37.6	6.7	592	CF795053	CF795053 891143 MA
14	37.4	6.6	554	AA975697	AA975697 EST187806
15	37.2	6.6	294	AA983391	AA983391 cq56e01.s
16	37.2	6.6	369	BG330743	BG330743 na39e01.
17	37.2	6.6	369	BUN94839	BUN94839 AGENCOURT
18	37.2	6.6	392	A1651148	A1651148 wa97h01.x
19	37.2	6.6	392	BES02665	BES02665 h220e10.x
20	37.2	6.6	395	BES02405	BES02405 hyl1f03.x
21	37.2	6.6	406	A1365427	A1365427 qz09a08.x
22	37.2	6.6	414	A1571898	A1571898 to20h12.x

C	23	37.2	6.6	417	3	BP424001	BP424001
	24	37.2	6.6	442	1	AW501328	AW501328
	25	37.2	6.6	443	1	AW015811	AW015811
	26	37.2	6.6	450	1	A1363365	A1363365
	27	37.2	6.6	452	1	A1123638	A1123638
	28	37.2	6.6	461	1	A1126856	A1126856
	29	37.2	6.6	477	1	AA733044	AA733044
	30	37.2	6.6	488	1	A1571759	A1571759
	31	37.2	6.6	489	2	BB328847	BB328847
	32	37.2	6.6	493	2	BB328847	BB328847
	33	37.2	6.6	506	2	BT196583	BT196583
	34	37.2	6.6	513	1	A1858617	A1858617
	35	37.2	6.6	527	1	A1697948	A1697948
C	36	37.2	6.6	557	1	AW952452	AW952452
C	37	37.2	6.6	564	5	BO671953	BO671953
C	38	37.2	6.6	569	3	BM719509	BM719509
	39	37.2	6.6	573	3	BM682828	BM682828
	40	37.2	6.6	585	1	AA706790	AA706790
	41	37.2	6.6	586	3	BP366808	BP366808
	42	37.2	6.6	589	5	BO787648	BO787648
	43	37.2	6.6	608	6	CA309122	CA309122
	44	37.2	6.6	610	3	BM987383	BM987383
	45	37.2	6.6	636	3	BM675573	BM675573

ALIGNMENTS

RESULT 1
AG133394
LOCUS AG133394 603 bp DNA linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-146A21.R, genomic survey sequence.
ACCESSION AG133394
VERSION AG133394.1 GI:16663072
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of library PTB
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 603)
REFERENCE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Ageo Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimbee@sc.riken.go.jp, URL:http://hgp.sec.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS
Sequencing: M13Rev
LIBRARY Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 603
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-146A21.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC library"

FEATURES
source
ORIGIN

Query Match	22.8%	Score 128.4;	DB 10;	Length 603;
Best Local Similarity	99.2%	Pred. No. 7.7e-23;		
Matches 129; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY 1 GCAGCCCTGCCTCAGAAAACAGAAAGACCGACACTACGGTACTACCCCCATGTG 60
DB 473 GCAGCCCTGCCTCAGAAAACAGAAAGACCGACACTACGGTACTACCCCCATGTG 532

Oy 61 GCTGGAGGTGAGGGAGCCCTCCTTGAGAGCGAGGCCAGGCCGTCAGAGTGGGTGAACGGCA 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 533 GCTGGAGGCGAGGGAGCCCTCCTTGAGAGCGAGGCCAGGCCGTCAGAGTGGGTGAACGGCA 592

Qy	121	GGGGTCTTGC	130
Db	593	GGGGTCTTGC	602

RESULT 2
A1680295

LOCUS	AI680295	512 bp	hikvna	linear	BS1	15-DEC-1995
DEFINITION	tw62a11.x1 NCI_CGAP_Ut3 Homo sapiens cDNA clone IMAGE:2264252 3' , mRNA sequence.					
ACCESSION	AI680295					

VERSION	AI680295.1	GI:489047
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
1 (pages 1 to 312)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL *Unpublished (1997)*

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmett-Buck, M.D., Ph.D.

CDNA Library Preparation: Lion Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-60 cell clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbip/image/image.html
Insert Length: 614 Std Error: 0.00
Seq primer: -400P from GABco
High quality sequence stop: 303.

FEATURES	Location/Qualifiers
source	1. .312

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2264252"
/tissue_type="poorly-differentiated endometrial
adenocarcinoma, 2 pooled tumors"
/lab_host="DH10B"
/clone_1fb="NCI CGAP U73"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.45 kb. Life Technologies catalog #:
11541-018"

```

ORIGIN

Query Match	7.3%	Score 41;	DB 1;	Length 312;
Best Local Similarity	54.2%	Pred. No. 8.4;		
Matches 83: Conservative	0:	Mismatches 70:	Indels 0:	Gaps 0:

251 CCCCCCAAGTTGCACATCTCAGTCTTTGCAAGGCTTGAGGCTTCCAGTCAAGGGTCCC 310

Oy ATCAAGTACAGGCCAGGTGCTGCAAGAAAGTCCCTGCAAGTCATGATAAAACAAAGGAGG 370
 Db AGGAACTCCACACAAAGAACTCTCAACAACGGCTGCGCTTCGACAGTCAGTCCAGAAACA 265
 Oy CTTGGGAAACCAATCTGAAGGCACTGCTTTG 403
 Db CTGACCATCCCTTGAAGGAGACCAAGCGCTTTG 302

RESULT 3
CNS0091P/C

DEFINITION
Drosophila melanogaster genome survey sequence TERT end of BAC # BACR19D16 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION	VERSION	KEYWORDS	SOURCE
AL053013	GI:4934461	GSS.	<i>Drosophila melanogaster</i> (fruit fly)

Eukaryota; metazoa; alcinopoda; hexapoda; annelata; freeygota
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999)
Genoscope - Centre National de Sequencage :

COMMENT

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazuhiro Oosagawa and Aaron Mammone in Pierer de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	Location/Qualifiers
source	1. .925

```
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_1ib="RRC1-98"
/note="end : TET3"
```

ORIGIN

Query Match	7.2%	Score 40.4;	DB 10;	Length 925;
Best Local Similarity	13.2%;	Pred. No. 15;		
Matches 40;	Conservative 135;	Mismatches 127;	Indels 0;	Gaps 0

QY 250 GCCCCCCCAA GGTTCACACTCA CTGCGCTT TGCAGGGTTGGGGCTTCCAGTTCACAGGGTCC 309
::: :::: :
Db 924 SBSCSCSCSCSSBSSSMSTSSNSBBSCSSSBSSBSSTSSMSSSSSBSSSSSSSSS 865

OY 310 CATCCACGTAACCAGCCCAGTGGCTGCAGAAGTCCTCGATCATGAACAACCAAGGAG 369
 :::|::|
D8 864 SGTSSAACVVCNASSSCTGCCCGCAGMABCMSSSSSCGSAAPGVVRASGAGAGRGGSG 805
 :::|::|
 :::|::|

[illegible]

Oy 430 GGGCAAGGCCACCGATGTGATGCAGCCAGAAGCAGGAAGCTGTCCCAGACTGCC 489
nb : : : : :
744 ..::..... : : : : :
...cccccccccvcvccccacaccaccccccaaaabccccccccaaaacagcgcctttt gaa

RESULT 4	CNS0091P	925 bp	DNA	linear	SSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION	BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL053013				
VERSION	AL053013.1	GI:4934461			
KEYWORDS	SSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 925)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
COMMENT	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
FEATURES	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mamonov in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw ap, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
FEATURES	Location/Qualifiers				
SOURCE	1..925				
	/organism="Drosophila melanogaster"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:7227"				
	/clone="BACR19D16"				
	/clone_1lb="RPCI-98"				
	/note="end : TET3"				
ORIGIN					
Query Match	7.1%; Score 40; DB 10; Length 925;				
Best Local Similarity	16.8%; Pred. No. 19; Mismatches 112; Indels 0; Gaps 0;				
Matches	45; Conservative 111; Mismatches 112; Indels 0; Gaps 0;				
Y	235 TGTGTGGCGGCATGCCCCCAAGTTGACACACTGACCTTGACAGGTTGGGCTT	294			
Db	554 TSSGSGVKGKSSGSGBSGCCSSSCSSSSCSBCCCCSCSVCSSBSBSKCSST	613			
Y	295 CCACTGACAGGCTCCCATTCACATGACGCCAGTGGCTGACAGAGTCCCTGCAGTC	354			
Db	614 SBSCSCCCKSSVCGTSCSSSSSCSSSSSTSSSSSTSKSSSGSSSSSYTTSK	673			
Y	355 ATGAACAACAGAGAGCTTGGGAAACACATCTGAAGGCATGCTTGATTGATGAGA	414			
Db	674 STASGSSSWAGGSSGSSSTSSSSSSSTSSSVSSGSKSTBSGSSSSGSSS	733			
Y	415 GGGTGGGCTTGAGCTGGGACCAAGGCTTGATGTCAGAGCCAGAGGACGAGACTG	474			

Db	734	SSTSSBBCTSSSSSSSSSSTCCCTCCCYSYSSSTSSSSSTSMGSTRSGSSSSSVG	793
Oy	475	GTCGCCAGCACTGCCCGCCTCTGC	502
Db	794	TSSSDSTSTCCSCCYWCCTCSFYBMB	821
RESULT 5			
LOCUS	BQ948345		
DEFINITION	BQ948345	915 bp	mRNA
ACCESSION	AGNCOURT_8803539 NIH_MGC_101 Homo sapiens	linear	EST 21-AUG-2002
VERSION	5', mRNA sequence.		
KEYWORDS	BQ948345		
SOURCE	BQ948345.1 GI:22363823		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
TITLE	Hominidae; Homo.		
JOURNAL	1 (bases 1 to 915)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@ds-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LCM2609 row: j column: 11 High quality sequence start: 24 High quality sequence stop: 233. Location/Qualifiers 1..915 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6426346" /tissue_type="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 101" /note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
FEATURES			
source			
Query Match	7.0%; Score 39.6; DB 5; Length 915;		
Best Local Similarity	51.7%; Pred. No. 24;		
Matches	90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;		
Oy	79	TCTGAGGCAAGGGCAGGGCAGCCGTGAGTGACGGCAGGGGTTCTTGCCATGTGG	138
Db	69	TCCGGCCCCAGAGGTGTGTCACAGTCGTGGGGGGTCACTGGCGACGCCAAGGCTTGGGGG	128
Oy	139	GCAACAGGGGTGCTATAAGCTTACTACGTACATCATCATGATCCCTTGTCGACCCTTGGA	198
Db	129	CCACCGAGGCCCACTTGCGTGTCTCCACCCGAGCATGCTTGGGCCCGGAGACTTAC	188
Oy	199	AGTCTGAAGTAGACAATTGTTCCATTAAAGAAAAGTGTGGCGGCGCATGCC	252
Db	189	AGTGAAGAAGGACATTGTTCTCTTGATGTGACCCCAAGCACGCCACGAAGCC	242
RESULT 6			

LOCUS	1101 bp	DNA	linear	GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL108460			
VERSION	AL108460.1 GI:5628764			
KEYWORDS	GSS.			
ORGANISM	Drosophila melanogaster (fruit fly)			
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peiryygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91060 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)			
COMMENT	- Web : www.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.			
FEATURES	Location/Qualifiers			
source	1..1101 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACN37L08" /clone_1lb="DrosBAC" /plasmid="pBelobAC11" /note="end : SP6"			
ORIGIN				
Query Match	7.0%; Score 39.6; DB 10; Length 1101;			
Best Local Similarity	14.0%; Pred. No. 25;			
Matches	43; Conservative 153; Mismatches 111; Indels 1; Gaps 1;			
QY	238 GTGCGCGCCATGCCCCCAACGTTGCACACTGACTGCTTTCAGCGGTTGGGCTTCCA 297			
DB	762 GAGGAGGGGGCCGACSSGVCGAACSSASCSASACGVSSSGSCGASCCGVSSCSAVAS 821			
QY	298 GTCAACAGGATCCCATCCACAGTACACGCCACGAGGTGCTCAGAAAGTCCCTGCAGTCA 357			
DB	822 SASVMSKVAAGVAVASCSAAGSMAGVSSSCRSSVASVAAAASVSSSSSSSVSS 881			
QY	358 AAACCAAGGAGAGGCTTGGGAAAACACATCTGAAGGGCATGCTTGAATTATAGTAGAGG 417			
DB	882 AAAVASASSSSASASMAVAVAVAVAVVSSVSSSSSCSSSSASAVVSVASVAAAS 941			
QY	418 TGGGGCTGGCTGGGCAAGCCACAG-GTCTGATCAAGAGCAGGACAGAACTGCT 476			
DB	942 SVSSSSSSSVTSSASVSVSAVMSAVSSSSSSSVSVVAVAASAAAAAAA 1001			
QY	477 CCCCAGCATGCGCCCGCTCTGGAATGAGTCTCTGCGCACCTGAGAACTGCT 536			
DB	1002 AAASSSSAAVAVVAASSSSSSSSSSSSSSSSVSSSCSVSVSSSVASVAVAS 1061			
QY	537 AGAGAGCC 544			
DB	1062 ASASASVS 1069			
RESULT 7				
LOCUS	AI289637 321 bp mRNA linear EST 21-DEC-1998			
DEFINITION	qW29d03.x1 NCI_GCAP_Ut4 Homo sapiens cDNA clone IMAGE:1992485 3',			
ACCESSION	AI289637			
KEYWORDS	mRNA sequence.			

VERSION	AI289637.1	GI:3932817
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
	Hominidae; Homo.	
REFERENCE	1 (bases 1 to 321)	
AUTHORS	NCI-CCAG http://www.ncbi.nlm.nih.gov/ncicgap.	
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	
JOURNAL	Tumor Gene Index	
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgaps@remail.nih.gov Tissue Procurement: Christopher Mookluk, M.D., Ph.D., Michael R.	
	Zimmer-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.llnll.gov/bdip/Image/Image.html Insert Length: 2780 Std Error: 0.00 Seq primer: -40UP from Gldco High quality sequence stop: 290.	
FEATURES	Location/Qualifiers	
Source	1..321	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:1992485"	
	/tissue_type="serous papillary carcinoma, high grade, 2	
	pooled tumors"	
	/lab_host="DH10B"	
	/clone_id="NCI CGAP Ut4"	
	/note="Organ: uterus; Vector: PCMV-SPOrt6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: oligo dt. Average insert size 1.48 kb. Life Technologies catalog #:	
	11542-016"	
ORIGIN		
Query Match	6.9%; Score 38.8; DB 1; Length 321;	
Best Local Similarity	57.4%; Pred. No. 32;	
Matches	70; Conservative 0; Mismatches 52; Indels 0; Gaps 0;	
Gy	251 CCCCCCAAGTTGCACACTCCTTCCTTGAGGGGTTCCTTCAGTCACAGGTTCCC 310	
Dd	171 CCGGCCAATGCTGCTCAGTTCCTCCCTCGGAGGAAGGGCCGAGGGGCTTCCTGTC 230	
Gy	311 ATCCACGTACCAGGCCGAGGTGGCTGTCAGAAAGTCCCTCCGAGTATGAAACCAAGGAGG 370	
Dd	231 AGGAACTCCACACAGAGGTGTGCAGACGGGCTGTGCTGCAGTCAGTCCACAGAACA 290	
Gy	371 CT 372	
Dd	291 CT 292	
RESULT 8		
CL542202	637 bp DNA linear GSS 14-JUN-2004	
LOCUS	OB_Ba0066B18.f OB_Ba Oryza brachyantha genomic clone	
DEFINITION	OB_Ba0066B18 5', genomic survey sequence.	
ACCESSION	CL542202	
VERSION	CL542202.1 GI:47612804	
KEYWORDS	GSS.	
SOURCE	Oryza brachyantha	
ORGANISM	Oryza brachyantha	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
	Erbatroidae; Oryzaceae; Oryza.	
REFERENCE	1 (bases 1 to 637)	
AUTHORS	Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,	

TITLE Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and Wing, R.
JOURNAL OMAR Project
COMMENT Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG CG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert length: 145000 Std Error: 0.00
Plate: 0066 row: B column: 18
Seq primer: TAA TAC GAC TCA CTA TAG CG
Class: BAC ends.

FEATURES
source Location/Qualifiers

1..637
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_Ba0066B18"
/tissue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_lib="OB_Ba"
/note="Vector: pAG1BAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 6.8%; Score 38.2; DB 10; Length 637;
Best Local Similarity 63.7%; Pred. No. 53;
Matches 58; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 411 GAGAGGCTGGGCTGGCTGGGCGCAAGCCACCAAGCTGTGATCAGACGACGAGCGAGGA 470
DB 462 GCGAAGATGGGGCGGAGCGGGCGAGGACGACCAAGGCGGCAACCCGCGGCGAGGA 521

QY 471 GCTGTCCCGACGACTGCGCCGCGGCTGTGC 501
DB 522 CCGAGGCCACGAGACTCCCGCGCGCGCGC 552

RESULT 9
BO931784 867 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8803536 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6426298
DEFINITION 5', mRNA sequence.
ACCESSION BO931784
VERSION BO931784.1 GI:22346815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LNC2609 row: h column: 11
High quality sequence atop: 627.
Location/Qualifiers
1..867

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6426298"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match 6.8%; Score 38.2; DB 5; Length 867;
Best Local Similarity 48.8%; Pred. No. 57;
Matches 103; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 79 TCCGAGGCGAGGCGGCGAGCCGTCAGGTGGTGACGCGAGGGCTTTCATGTGG 138
DB 70 TCCGCGCCCGAGGTGGTGGCCGCTCGGGGGGTCACTGGCGGAGCCAGGCCCTGGGGG 129
QY 139 GCACAGGGGCTGCATACAGCTTACCTACAGTCAATCGAGTCCCTGGTGCAGCCCTTGGGA 198
DB 130 CCACCGAGGCCCACTCGCTGTGTCACCCGAGACATCGCTTGGGCCCCGCACTAC 189
QY 199 AGTCGTGAAGTGAAGCAATGTTCCATTAGAAAGTGTGTGGCCGCGCATGCCCCCA 258
DB 190 AGTGAAGGAGGAGGACCTGTGTTCTCCTTGATGAGCAACGAGGAGCCAGAGCCGATGGC 249
QY 259 CGTGCACACTCATCTGCTTTGCAAGGGTTGG 289
DB 250 CATGACAAAGGCGCGGTGATGATGAGCAGG 280

RESULT 10
A1970105 393 bp mRNA linear EST 20-OCT-2000
LOCUS w989c02.x1 NCI_CGAP_G66 Homo sapiens cDNA clone IMAGE:2479202.3',
DEFINITION mRNA sequence.
ACCESSION A1970105
VERSION A1970105.1 GI:5766931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 393)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Patricia
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 488 Std Error: 0.00
Seq primer: -40UP from Glbco.
Location/Qualifiers
1..393

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
 /clone="IMAGE:2479202"
 /issue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_1lb="NCI_CGAP GC6"
 /note="Vector: pT773D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
 from the normalized library NCI CGAP GC4 was prepared, and
 as circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (cloneIDb
 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 6.7%; Score 38; DB 1; Length 393;
 Best Local Similarity 54.2%; Pred. No. 55;
 Matches 77; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 251 CCCCCAAGTTCACACTCCTTGGCAGGGGCTTCCAGTCACAGGGTCCC 310
 |||||
 DB 150 CCGCCCATCTGGCTCAGTTCCTGCCCCCTGGAGGAGGAGGCGGCGCTTCTCTCCC 209
 |||||
 QY 311 ATCCAGTACCAAGCCAGGTGGCTGCAGAGGTCCCTCGCATGTAACCAAGGAGG 370
 |||||
 DB 210 AGGGAATCTCCACACAGAGGTCTGCAGACGGGCTGGCAGTCAATCCAGAAAC 269
 |||||
 QY 371 CTGGGAACACACATCTGAAGG 392
 |||||
 DB 270 ACTGACCATCTCCCTTAGGAGG 291
 |||||

RESULT 11 481 bp mRNA linear EST 03-MAR-2000
 AM519131 has0f09.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2877161 3',
 LOCUS mRNA sequence.

DEFINITION AM519131
 AM519131 GI:7157213

ACCESSION AM519131
 VERSION AM519131.1 GI:7157213
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

1 (bases 1 to 481)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

REFERENCE
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 COMMENT Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaps-r@mail.nih.gov

Life Technologies catalog #: 11548-013

Clone Sequencing by: Washington University Genome Sequencing Center
 DNA distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:

image.llnl.gov/image/html/resources.shtml
 image.llnl.gov/image/html/resources.shtml

Seq primer: -40UP from Gibco
 High quality sequence stop: 419.

FEATURES

source

1..481
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2877161"
 /issue_type="adenocarcinoma"
 /lab_host="DH10B"
 /clone_1lb="NCI_CGAP Pan1"
 /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.72 kb. Life Technologies catalog #:
 11548-013"

ORIGIN

Query Match 6.7%; Score 37.8; DB 1; Length 481;
 Best Local Similarity 56.6%; Pred. No. 64;
 Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 251 CCCCCAAGTTCACACTCCTTGGCAGGGGCTTCCAGTCACAGGGTCCC 310
 |||||
 DB 157 CCGCCCATCTGGCTCAGTTCCTGCCCCCTGGAGGAGGAGGCGGCGCTTCTCTCCC 216
 |||||
 QY 311 ATCCAGTACCAAGCCAGGTGGCTGCAGAGGTCCCTCGCATGTAACCAAGGAGG 370
 |||||
 DB 217 AGGGAATCTCCACACAGAGGTCTGCAGACGGGCTGGCAGTCAATCCAGAAAC 276
 |||||
 QY 371 CT 372
 |||||
 DB 277 CT 278
 |||||

RESULT 12

A1818367 523 bp mRNA linear EST 07-MAR-2000
 A1818367 wk59h08.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2419743 3',
 LOCUS mRNA sequence.

DEFINITION A1818367
 A1818367 GI:5437446
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

1 (bases 1 to 523)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

REFERENCE
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 COMMENT Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaps-r@mail.nih.gov

Life Technologies catalog #: 11548-013

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:

www.bio.llnl.gov/bbrp/image/html
 Insert length: 1852 Std Error: 0.00

Seq primer: -40UP from Gibco
 High quality sequence stop: 428.

Location/Qualifiers

FEATURES

source

1..523
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2419743"
 /issue_type="adenocarcinoma"
 /lab_host="DH10B"
 /clone_1lb="NCI_CGAP Pan1"
 /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.72 kb. Life Technologies catalog #:
 11548-013"

ORIGIN

Query Match 6.7%; Score 37.8; DB 1; Length 523;
 Best Local Similarity 56.6%; Pred. No. 65;
 Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 251 CCCCCAAGTTCACACTCCTTGGCAGGGGCTTCCAGTCACAGGGTCCC 310
 |||||
 DB 148 CCGCCCATCTGGCTCAGTTCCTGCCCCCTGGAGGAGGAGGCGGCGCTTCTCTCCC 207
 |||||
 QY 311 ATCCAGTACCAAGCCAGGTGGCTGCAGAGGTCCCTCGCATGTAACCAAGGAGG 370
 |||||
 DB 208 AGGGAATCTCCACACAGAGGTCTGCAGACGGGCTGGCAGTCAATCCAGAAAC 267
 |||||

QY 371 CT 372
||
Db 268 CT 269

RESULT 13
LOCUS CF795053/c 592 bp mRNA linear EST 21-OCT-2003
DEFINITION 891143 MARC 4P1G Sue scrofa cDNA 3', mRNA sequence.
ACCESSION CF795053
VERSION CF795053.1 GI:37799626
KEYWORDS EST.
SOURCE Sue scrofa (pig)
ORGANISM Sue scrofa
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.

REFERENCE 1 (bases 1 to 592)
AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonnenman,D.J., Wray,J.B. and Keefe,J.W.
TITLE Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smtth@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
plate: TMM8015 row: K column: 23
Seq primer: TAGAAGCAGCTGAGG.
Location/Qualifiers

FEATURES
source 1..592
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/cissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4P1G"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match 6.7%; Score 37.6; DB 6; Length 592;
Best Local Similarity 50.6%; Pred. No. 76;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 353 TCATGAAACCAAGAGGCTTGGAAACACATCTGAAGGCGATGCTTTGATTAGTGA 412
|||
Db 234 TCAGAAACCAAGAGGAGACTGATTAATAATACGAGGGGGCTGTTTATGATGTCA 175
|||
QY 413 GAGGCTGGGCTGGGCTGGGCAAGGCCACAGCTGTGATGACAGCCAGAGCGAAGC 472
|||
Db 174 CCGGGCCCACTACTCTGGGAGGCTGTGAGGTGTGCGGCTACGCCCTGGCCAGCTGTG 115
|||
QY 473 TGGTCCCGAGACTGCCCGCGCTCTGCGATGACAGTCTCTGGCCACTGAGAACAGC 532
|||
Db 114 CCGTCCAGGGCTCGCGCTTCCGCTTACCTTGTGCTGTGACACAGGCTCAGC 55
|||

RESULT 14
LOCUS AW975697/c 554 bp mRNA linear EST 02-JUN-2000
DEFINITION EST387806 MAGE resequencees, MAGN Homo sapiens cDNA, mRNA sequence.
ACCESSION AW975697
VERSION AW975697.1 GI:8166915
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 554)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gasparid,R., Gay,C.,
Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: john@igir.org
Plate: 352
Seq primer: Forward.
Location/Qualifiers

FEATURES
source 1..554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequencees, MAGN"
/note="Vector: pBluescriptSKm"

ORIGIN
Query Match 6.6%; Score 37.4; DB 1; Length 554;
Best Local Similarity 55.9%; Pred. No. 84;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 43 GTGACTACCCCTGCTGCTGCTGAGGTGAGGAGCTCTGAGGCGAGGCGAGCC 102
|||
Db 545 GTTCTTAAGCCCTTAATTTGCAAGGTGTGAACCCCTATGGGCGCTGGCCAAATCAGCA 486
|||
QY 103 GTGAGTGGGTGACGAGGCGGCTTGCATGTGTGGGACACAGGGCTGCATACGCTTAC 162
|||
Db 485 AGAAGGGGCTGAAGAAAGGCGACGCTTCTTGGGCTTCACTGGGACTGGAAGAGCTGCC 426
|||
QY 163 TCAGTGA 169
|||
Db 425 TTCCTGA 419
|||

RESULT 15
LOCUS AA983291 294 bp mRNA linear EST 23-JUL-1998
DEFINITION o656e01.81 NCI_CGAP_kids Homo sapiens cDNA clone IMAGE:1590360 3',
mRNA sequence.
ACCESSION AA983291
VERSION AA983291.1 GI:3161816
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 294)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.1lnl.gov/bdirp/image/image.html

Insert Length: 890 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 273.
Location/Qualifiers

FEATURES
source

1.294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1590360"
/issue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_11b="NCI CGAP_K1d5"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AACTGAGAGATTCCGGCGCAATATTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 6.6%; Score 37.2; DB 1; Length 294;
Best Local Similarity 56.6%; Pred. No. 84;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 251 CCCCCAAGTTGACACTCACTGCTTGGCAGAGTTGGGCTTCCAGTCACAGGGTCCC 310
Db 159 CCGCCCATGCTGCTCAGTTCCTGCCCCCTGGAGGAGAGGGCCGAGGGGCTTCCGTC 218
QY 311 ATCCACGTACCAAGCCGAGGTGGCTGCGAAGGTCCCTGCGAGTCATGAACCAAGGAGG 370
Db 219 AGGGAAGTCCACACAGGAGGTCTGCAGACGGGCTGGCTGCGAGTCAGTCCAGGAA 278
QY 371 CT 372
Db 279 CT 280

Search completed: March 25, 2006, 18:16:22
Job time : 3245 secs


```

: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Wiegand, Royer C.
: TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-10(115849) B
: CURRENT APPLICATION NUMBER: US/09/902,540
: CURRENT FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: 60/217, 883
: PRIOR FILING DATE: 2000-07-10
: NUMBER OF SEQ ID NOS: 16825
: SEQ ID NO 604
: LENGTH: 4050
: TYPE: DNA
: ORGANISM: Myxococcus xanthus
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(4050)
: OTHER INFORMATION: unsure at all n locations
US-09-902-540-604

```

Query Match	6.2%;	Score 35.2;	DB 3;	Length 4050;
Best Local Similarity	53.7%;	Pred. No. 5.5;		
Matches 73; Conservative	0;	Mismatches 63;	Indels 0	

Query Match	6.2%	Score 35.2;	DB 3;	Length 4050;
Best Local Similarity	53.7%;	Pred. No. 5.5;		
Matches 73; Conservative	0;	Mismatches 63;	Indels 0;	Gaps 0;

407 TAGTGAGAGAGGCTGGGGCTGGGCTGGGCGCAGGGCCACAGGCTTGAGTCAAGAGCCGAGGCA 466
3587 TCGTGGCGCGTGC CGGCTTCGAGGAGCCCCACAGACACGTCCTTGCGCGCGTCCCTGGGC 3648

467 GGAAGCTGGTCCCGACGACCTGGCCGCGCTGTGGATGAGTCTCTGGCCACTGAG 526
3647 CGGCTCCCTTGACGGAACTCATACCGGACGGCCATGAGTCCCCCTGGGCTTCTTG 3706

QY 527 AACAGCCTGTAGAGAG 542
Db 3707 AGCTGCCAGTAGCGAG 3722

RESULT 6
US-09-949-016-12065

Sequence 12065, Application US/09949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ATTORNEY: STEVENSON STRONG & LEE

```

? FILE REFERENCE: C:\D0U3107
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 12065
? LENGTH: 108440

```

```

; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12065

```

Query Match	6.2%;	Score 35.2;	DB 3;	Length 108440
-------------	-------	-------------	-------	---------------

Best Local Similarity 58.7%; Pred. NO. 16;
Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Oy 66 AGGTGAGGAGACCTCCTGAGGAGGAGGAGGAGGAGCGGTCAAGTGGGTGAACGGACGGGGT 125
 |||||
 Db 81526 AGGTGTTAAACCACTATGGGGCTTGCCAAATGACAAAGAGGGGCTTGAGAGGAGGGCA 81585

126 CTTGCCATGTTGGGACAGGGCTGCATACAGCTTACTCAGTGA 165

Db 81586 GCTGTCCTGGGGTACACTGGACTGGAAGAGCTGCTTCCCTGA 81629

RESULT 7
US-09-949-016-14090
; Sequence 14090, Application US/09949016
; Patent No. 6817328

; FALCONE NO.: 0042537
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: VENTNER, J. Craig et al.
 ;
 ; TITLE OR INVENTION: POLYMERIZING BY VARIOUS GROUPS 10000047MM

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT PRIORITY DATE: 04/04/00

;
; CURRENT FILING DATE: 2000-04-14
;
; PRIOR APPLICATION NUMBER: 60/241,755
;
; PRIOR FILING DATE: 2000-10-20
;
; PRIOR APPLICATION NUMBER: 60/237,768

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14090
; LENGTH: 108441
; TYPE: DNA

```

ORGANISM: Human
US-09-949-016-14090

Query Match	6.2%	Score 35.2	DB 3	Length 108441
Best Local Similarity	58.7%	Pred. No. 16		
Matches 61; Conservative	0	Mismatches 43	Indels 0	

66 AGGTGAGGGAGCCTCTTGAGGCGAGGCGCAGGGCAGCCGTCAAGTGGTGACCGGACGGGGT 125

Db 81526 AGGTGGTGAACCACTATGGGCTGGCCAAATCAGCAAGAGGGGGCTGGAGGAGGGCA 81581

Dy 126 CTTCGCATGCTGGGCACAGGGGCTGCATCACGCTTACTCAGTGA 169
||| ||| ||| ||| ||| ||| |||
Db 81586 GCTGTCTTTGGGCTACACTGGGACTGGAAGAAGCTGCCCTTCTGA 81629

RESULT 8
US-09-843-905A-11
; Sequence 11, Application US/09843905A

; Patent No. 6703487
; GENERAL INFORMATION:
; INFORMATION:

; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HYBRID DERIVATIVE POLYMERIZATION

FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: IIS/09/A43 905A

/ CURRENT FILING DATE: 2001-04-27
 / PRIOR APPLICATION NUMBER: US 60/200,198

; PRIOR FILING DATE: 2000-04-28
 ; NUMBER OF SEQ ID NOS: 15

```
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 11
```

```

; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: m1ac feature

```

NAME/DOB: [REDACTED] / [REDACTED]
LOCATION: (513) .. (513)
OTHER INFORMATION: Unsub

US-09-843-905A-11

Query Match 6.24; Score 35; DB 3; Length 1338

Best Local Similarity 49.2%; Pred. No. 4.4;
Matches 92; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 364 AGGAGGCTTTGGGAAACCAATCTGAAAGGCATGCGTTGATTTAGTGAAGAGGCTGGGGC 422

Db 794 AGGAGGCGCTCTTCATCGACTGTGTGGGSCACAACTGCTGTGGCCACACCGGCGGGGC 852


```
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15483
/ LENGTH: 6872
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(6872)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15483
```

Query Match 6.0%; Score 33.8; DB 3; Length 6872;
Best Local Similarity 54.4%; Pred. No. 16;

Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Oy 330 AGGCGATGCTTTGATTAGTACAGAGGCTGCGGCTGGGAGGACACAGATCTG 449
Db 1635 AGGATGGGCTGAGTTAAGTACGAAAGCTGCCAGTCCCGAGAAAGTCTTCCCAAG 1566
Oy 450 AGTCAGAGCCAGAGGACGAGAACTGTCTCCAGCACTGCCCGCCCTCTGCATGACAT 509
Db 1565 AACGAGGCAAGGTCAGCTTGGGCGCACTTACCTGCTCTTGAAGGATGTGT 1506
Oy 510 CCTCC 514
Db 1505 CTCCC 1501
```

RESULT 13

```
US-09-949-016-15980/C
/ Sequence 15980, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15980
/ LENGTH: 239527
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(239527)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15980
```

Query Match 6.0%; Score 33.8; DB 3; Length 239527;

```
Best Local Similarity 48.7%; Pred. No. 52;
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
Oy 336 CAGGTGCTGAGAAAGTCCCTCGCATGATGAACCAAGGAGGCTTGGAAACCATAT 385
Db 194917 CAGGAGACTTAAAGTCAAGCAGACAGACAGTCAACCATATGCTGTAGCCCAAC 194858
Oy 386 CTGAAGGGCATGCTTTATTAGTGAAGGTGGGCTGGGCTGGGCAAGCCCAAG 445
Db 194857 AGACACTCAACACCAAGGCTGGGTGAGCTGTGTGTGGCAATCTCTGTGCTGA 194798
Oy 446 TCTGAGTCAGAGCCAGAGGACGAGAACTGTCTCCAGCACTGCCCGCCCTCTGCATG 505
Db 194797 CTCCACTGGAGAGATGCTGTGGCACTTGAACCAAGCACTCCCTGAGACCTGCTATG 194738
Oy 506 CAGTCTCC 514
Db 194737 CAGTCTTC 194729
```

RESULT 14

```
US-09-621-976-17202/C
/ Sequence 17202, Application US/09621976
/ Patent No. 6639063
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Jobert, S.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: ESTS and Encoded Human Proteins.
/ FILE REFERENCE: GENSET.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621,976
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 17202
/ LENGTH: 364
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-621-976-17202
```

Query Match 6.0%; Score 33.6; DB 3; Length 364;
Best Local Similarity 10.9%; Pred. No. 7.1;

Matches 30; Conservative 126; Mismatches 120; Indels 0; Gaps 0;

```
Oy 9 GCTTCAGAAAACAGAGAGCGACACACTCAGGTGATCACCCTCATGTGGTGAAG 68
Db 292 SCWYSRKQMSCMYMTGRMSMMKSTYKRMBSRBMWTGYRMSMMWTGSTRCSKKR 233
Oy 69 TGAAGGAGCCCTCTGAGGAGGCGGAGGCGGAGCGCTCAGGTGGGTGACGAGGGTCTT 128
Db 232 KKGSTSSKYASTSGSSSKYMTCKRSKRYSATYYSCMMKWKYCMMSATYSGCMMW 173
Oy 129 GCCATGTGGCAGACAGAGGCTGATACAGCTTACTAGTACATGACATGCTGCTGCC 188
Db 172 RMYVSCMSRYSVSTYSRKGCSCTGKWKGCYRYRMYRGRMYRMRKARRTKTSRG 113
Oy 189 AGCTTGTGAGTCTGAAAGTGAAGCAATGTTCCATTAAGAAAGTGTGTGGCCGCA 248
Db 112 WCRSTKRYRTCAWGAAGTGMCCMRMGSTGASVYRKMSKMSKMSKMGYVSYMSYM 53
Oy 249 TGCCCCCAACGTTGACACACTGCTGCTTGGCAG 284
Db 52 SSGSGSWKSCGSGSSYYCRSYSSWGKSRCTCCWG 17
```

RESULT 15

```
US-09-949-016-94858
/ Sequence 94858, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: ParseSeq for Windows Version 4.0
SEQ ID NO 94858
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-94858

Query Match 6.0%; Score 33.6; DB 3; Length 601;
Best Local Similarity 50.6%; Pred. No. 8.4;
Matches 78; Conservative 1; Mismatches 75; Indels 0; Gaps 0;

QY	346	CTGCGAGTCATGAAACCAAGGAGGCTTGGAAACCATCTGAAGGCGATGAGCTTTGAT	405
DB	149	CTGGTAAGCCAGGCGCCACAGGGTGTGAGAGACGCTCAACACTGGGGCGGGCTGCGG	208
QY	406	TTAGTGAGAGGCTGGGCTTGGCTGGGCAAGGCCACAGGTCCTGAGTCAGAGCCAGAGGC	465
DB	209	TGAGTAGCAGGGAATGGGCTGGGCAAGCTGGGGACCCCTGGGGGCTCTCCATGAGACAAGA	268
QY	466	AGGAAGCTGTCCCGCAGCACTGCCCGCCGCTCT	499
DB	269	AGGAACAGCATCTGAGAGACCTGACCAACCCCTCT	302

Search completed: March 25, 2006, 15:02:01
Job time : 202.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2006, 16:12:32 ; Search time 3244 Seconds
(without alignments)
8134.376 Million cell updates/sec

Title: SEQ1-69C
Perfect score: 564
Sequence: 1 gcagccctgcctcagaaac.....agtcgcctcttcgacttc 564

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_eest1:
2: gb_eest2:
3: gb_eest3:
4: gb_hrc:
5: gb_eest4:
6: gb_eest5:
7: gb_eest6:
8: gb_eest7:
9: gb_gsest1:
10: gb_gsest2:
11: gb_gsest3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	130	23.0	603	AG133394	AG133394 Pan trogl
2	41	7.3	312	A1680295	A1680295 tw62a1.x
3	40.4	7.2	925	CNS0091P	AL053013 Drosophi1
4	40	7.1	925	CNS0091P	AL053013 Drosophi1
5	39.6	7.0	915	BQ948345	BQ948345 AGENCOURT
6	39.6	7.0	1101	CNS0178Y	AL108460 Drosophi1
7	38.8	6.9	321	A1289637	AL1289637 gw29d03.x
8	38.2	6.8	637	CL542202	CL542202 OB_Ba006
9	38.2	6.8	867	BQ931784	BQ931784 AGENCOURT
10	38	6.7	393	A1970105	A1970105 wg89c02.x
11	37.8	6.7	481	A1818367	A1818367 wa50f09.x
12	37.8	6.7	523	A1818367	A1818367 wa50f09.x
13	37.6	6.7	592	CF795053	CF795053 891143 MA
14	37.2	6.6	364	AA983291	AA983291 cg56e01.s
15	37.2	6.6	364	BG230743	BG230743 natf39e01
16	37.2	6.6	369	BUI94839	BUI94839 AGENCOURT
17	37.2	6.6	392	A1651148	A1651148 wa97h01.x
18	37.2	6.6	392	BES02665	BES02665 hz20e10.x
19	37.2	6.6	395	BES02405	BES02405 hy12f03.x
20	37.2	6.6	406	A1365427	A1365427 qz09a08.x
21	37.2	6.6	414	A1571898	A1571898 to20h12.x
22	37.2	6.6	417	BP424001	BP424001 BP424001

23	37.2	6.6	442	1	AW501328
24	37.2	6.6	443	1	AW015811
25	37.2	6.6	450	1	A1363365
26	37.2	6.6	452	1	A1123638
27	37.2	6.6	461	1	A1126856
28	37.2	6.6	477	1	AA733044
29	37.2	6.6	488	1	A1571759
30	37.2	6.6	489	2	BE328847
31	37.2	6.6	493	1	A1096775
32	37.2	6.6	506	2	BF196583
33	37.2	6.6	513	1	A1858617
34	37.2	6.6	527	1	A1697948
35	37.2	6.6	557	1	AW952452
36	37.2	6.6	564	5	BQ671953
37	37.2	6.6	569	3	BM719509
38	37.2	6.6	573	3	BM682828
39	37.2	6.6	585	1	AA706790
40	37.2	6.6	586	3	BP366808
41	37.2	6.6	589	5	BQ787648
42	37.2	6.6	608	6	CA309122
43	37.2	6.6	610	3	BM987383
44	37.2	6.6	636	3	BM675573
45	37.2	6.6	643	5	BU947551

ALIGNMENTS

RESULT 1
AG133394
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of library PTB
Unpublished
2 (bases 1 to 603)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-ANG-2001) Ageo Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpanzee@ac.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 603
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-146A21.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

[illegible]

QY	311	ATCCAGCTACGAGCCGAGGCTGGTGGCAAGGTCCTCCGAGTCATGAACCAAGGAGG	370
Db	210	AGGGAATCTCCACACCAAGAGTCTGCAGACGGCTGGCCCTCGAGTCAGTCCAGGAACA	269
QY	371	CTTGGGAAACCATCATCTGAAGGGCATGCGCTTTG	403
Db	270	CTGACATCTCCCTTAGGAGAGACACGCGCTTG	302
RESULT 3			
CNS0091P/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
SOURCE			
ORIGIN			
Query Match	7.2%;	Score 40.4;	DB 10; Length 925;
Best Local Similarity	13.2%;	Pred. No. 15;	
Matches	40;	Conservative 135;	Mismatches 127; Indels 0; Gaps 0;
QY	250	GGCCCCCAACGTTGACACATCATCTGCGCTTTGAGGGCTTCCAGTCACAGGCTCC	309
Db	924	SBGSCGCCGCSBSCSSSMSTSSNSBSCSSBSSSSSTSSMSSSBSSSSGSSSSS	865
QY	310	CATCCACGTAACCAAGCCAGGTGCTGTGACGAAGTCCCTCGCATCATGAACAAGGAG	369
Db	864	SGTSSACVKNAASSCGCCCGMABCCMGSSSSSCGSASARGVKVASGAGKRGGS	805
QY	370	GCTTGGGAAACCATCATCTGAAGGCAATGCGCTTTATTTAGTAGAGGCTGGCGCTGG	429
Db	804	GASASHSSSSSCBSSSSSCSASCMASASSSSASSSRSRSGGAGSGASASSSSSSA	745
QY	430	GGGCAAGCCACCAAGCTAGTGTAGAGACCAAGGACGAAGAACTGATCCCAAGCATG	489
Db	744	SAGSVSSASSSSSCSSSVSCSSVASMSGCSBSSSSASBSSSSSSASCAACSCCTT	685

OY		490	CGCCGCCTCTGTGGAGTGAAGTCCTCCTGCGCACCTTGAGAAGACGCCCTGTAGAGCGCAGTGG	543
Db		664	SWSGCSCTASMSAARSSSSSSSCSSSSSMASASSSSASSSSSSGSAGCBBS	625
OY		550	CG 551	
Db		624	MS 623	
RESULT 4				
LOCUS	CNS0091P			
DEFINITION	CNS0091P 925 bp DNA linear GSS 03-JUN-1999 Drosophila melanogaster genome survey sequence TET3 end of BAC # BACH19D16 of RPCT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL053013			
VERSION	AT053013.1 GI:4934461			
KEYWORDS	GSS.			
SOURCE	Drosophila melanogaster (fruit fly)			
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peirysgota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 925) Genoscope.			
REFERENCE	Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)			
AUTHORS	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.			
TITLE	Location/Qualifiers			
JOURNAL	1..925			
COMMENT	/organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACH19D16" /clone_lib="RPCT-98" /note="end : TET3"			
FEATURES				
Source				
ORIGIN				
Query Match	7.1%; Score 40; DB 10; Length 925;			
Best Local Similarity	16.8%; Pred. No. 19; Mismatch 112; Indels 0; Gaps 0;			
Matches	45; Conservative 111; Mismatch 112; Indels 0; Gaps 0;			
OY	235 TGTTGTGGCGGCATGCCCCCAAGTTCACACTCATCTGCTTTGACAGGTTGGAGCTT	294		
Db	554 TSSGSGVKGKGSGBSGBSCSCSSCSSCSBCBCCCSCS9SYCCSSBSBSKSST	613		
OY	295 CCAATCACAGGTCCTCCATTCACATACACGCCAGTAGCTGTACAAAGATCCCTGCACATC	354		
Db	614 SBSGSCCCSKSVCTGTCSSSSSCSSSSSTSSSSTKSSBSGSSSSSSSYTSX	673		
OY	355 ATGAACAACAAAGAGAGCTTTGGAAAACACATCTGAAGGACATGGCTTGATTATTAAGTAGA	414		
Db	674 STASAGSSSWACGSSGSTSSSSSSSTSSSVTS9SGSKSTBS9SBSSGSSSS	733		
OY	415 GGTTGGGGCTGGGCTGGGACAGCCACCAGGTCTGAGTCAGAGCCAAGACAGAGAGCTG	474		

Db	734	SSTSSBBCTSTSSSSSSSSSSSSSSCTCCSSYSSTSSSSSTSSSGTSGSSSSSVG	793
Qy	475	GTCCCCAGACATGCCCCGCCCTCTGCG	502
Db	794	TSSSDSTSTCCSCCTCSTYBMB	821

RESULT 5

LOCUS	BQ948345	915 bp	mRNA	linear	EST 21-AUG-2000
DEFINITION	AGENCOURT 8803539 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6426346				
ACCESSION	BQ948345				5', mRNA sequence.
VERSION	BQ948345.1	GI:22363823			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 915)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCM2609 row: j column: 11 High quality sequence start: 24 High quality sequence stop: 233. Location/Qualifiers 1..915 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6426346" /tissue_type="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_id="NIH MGC 101" /note="Organ: lung; Vector: pOTB7, Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."				

ORIGIN

Query Match	7.0%;	Score 39.6;	DB 5;	Length 915;
Best Local Similarity	51.7%;	Pred. No. 24;		
Matches	90;	Conservative	0;	Mismatches 84; Indels 0; Gaps 0;

Qy	79	TCCTGAGGCGAGGGCCAGGGCAGCGTCAAGTGGGTGACGGGAGGGGCTTGGCATGTGG	138
Db	69	TCCGGCCCCAGAGGTGTGGCCACGTCGTGGGGGGTCACTGGGCGAGCCAAAGGCTTGGGGG	128
Qy	139	GCAAGAGGGGTGATACAGTTACTCACTGATGACATGACAGTCCCTGGTGGCCAGCCTTGGGA	198
Db	129	CCACCGGAGGCCCACTCTGGGTGTCCACCCGGAGCATGTGTTGGGCCCGGCACTCAC	188
Qy	199	AGTCTGAAGTGAAGCATGTTTCCCATTAAGAAAGTGTGNGCGGCGCATGCC	252
Db	189	AGTAGCAAGAGGACATTGTTCTCTTGAATGACCCCAAGAGGCCCAAGAGCC	242

RESULT 6

LOCUS	CNS0175Y	1101 bp	DNA	linear	GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL108460				
VERSION	AL108460.1 GI:5628764				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101) Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
JOURNAL	- Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (BDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.				
FEATURES	Location/Qualifiers				
source	1..1101 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone_1fb="BACN37L08" /clone_1lb="DrosBAC" /plasmid="pBeloBAC11" /note="end : SP6"				
ORIGIN					
Query Match	7.0%; Score 39.6; DB 10; Length 1101;				
Best Local Similarity	14.0%; Pred. No. 25;				
Matches	43; Conservative 153; Mismatches 111; Indels 1; Gaps 1;				
OY	238 GTGCGCGGCATGCCCCCAACGTTGCACACTGCTTGTGACAGGTTGGGGCTTCCA 297				
Db	762 GGGGGRGGGGCCAGGVCAGAACSCASCSBAGMCGVSGSSCSASGSCCGVSSCSAVAS 821				
OY	298 GTCAAGAGGTCCCACTCCACGTCACGACCGCAGCGTGTGTGAGAAGGTCCTCGCACTCATG 357				
Db	822 SASVMSKVASVVAASCAVASGMSAGVASSCRSSVASSVSAASVSSVSSSSSSSVAS 881				
OY	358 AAACCAAGAGAGCTTGGGAAACCAACATCTGAAGGCACTGCTTGATTGATTAGTAGAGGG 417				
Db	882 AAASVASSSSASASMAVAAMAAVAAVVSVSVSSSSSSSSSSASVSVASVABAS 941				
OY	418 TGGGCTGTGGCTGGGCAAGCCACAG-GTCTGAGTCAGAGCCAGAGCAGAGACTGGT 476				
Db	942 SVSSSSSSSVSTSSASVSVSAVEMSAVSSASSSSSSSVSVVAASAAAAAAA 1001				
OY	477 CCCGACGACTGCCCGCCGCTCTGGGATGCACTCTCTCGGCGACACTAGAACACCTGT 536				
Db	1002 AAASSSSAAVAVVAASSSSSSSSSSSSSSVSSSVSSSVSSSVSSSVAVAS 1061				
OY	537 AGAGAGGC 544				
Db	1062 ASASASVS 1069				
RESULT 7	AI289637 321 bp mRNA linear EST 21-DEC-1999				
LOCUS	AI289637				
DEFINITION	QW29d03.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:1992485 3', mRNA sequence.				
ACCESSION	AI289637				

VERSION	AI289637.1	GI:3932817
KEYWORDS	EST.	Homo sapiens (human)
SOURCE	Homo sapiens	
ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 321)	
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/MLN.L at: www-bio.ln1l.gov/bbip/image/image.html Insert Length: 2780 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 290. Location/Qualifiers 1..321 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:1992485" /tissue_type="serous papillary carcinoma, high grade, 2 pooled tumors" /lab_host="DH10B" /clone_id="NCI CGAP Uc4" /note="Organ: uterus; Vector: pCMV-Sport6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.48 Kb. Life Technologies catalog #: 11542-016"	
ORIGIN		
Query Match	6.9%;	Score 38.8; DB 1; Length 321;
Best Local Similarity	57.4%;	Pred. No. 32;
Matches	70; Conservative	0; Mismatches 52; Indels 0; Gaps 0
OY	251	CCCCCAAGTTGACACTGCTGCGAGGTTGGGCTTCACATCAGGGTCCC 310
Db	171	CGGCCAAGCTGGCTCAATTCTGCTCCGTGAGGAGGGCCAGGGGCTTCGTGCC 230
OY	311	ATCCACGTACGAGCCGAGGTGGCTGACAGAGTCCCTCGCAGTCATGAAACCAAGGAGG 370
Db	221	AGGGAATCCACACACAGAGTGTGACAGCGGGCTGTGCTGCAGTCAGTCCAGGAACA 290
OY	371	CT 372
Db	291	CT 292
RESULT 8		
CL542202	637 bp	DNA linear GSS 14-JUN-2004
LOCUS	OB_Ba0066B18.f OB_Ba	Oryza brachyantha genomic clone
DEFINITION	OB_Ba0066B18 5', genomic survey sequence.	
ACCESSION	CL542202	
VERSION	CL542202.1	GI:47612804
KEYWORDS	GSS.	
SOURCE	Oryza brachyantha	
ORGANISM	Oryza brachyantha	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.	
REFERENCE	1 (bases 1 to 637)	
AUTHORS	Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,	

TITLE Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
JOURNAL OMAP Project
COMMENT Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert Length: 14500 Std Error: 0.00
Plate: 0066 row: B column: 18
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES

Location/Qualifiers
1..637
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_Ba0066B18"
/issue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_id="OB_Ba"
/note="Vector: pAG1BAC1, Site_1: HindIII, Site_2: HindIII"

ORIGIN

Query Match 6.8%; Score 38.2; DB 10; Length 637;
Best Local Similarity 63.7%; Pred. No. 53;
Matches 58; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Oy 411 GAGAGGCTGGGCTGGCTGGGCAAGCCACCAAGTCTGAGTCAGAGCCAGAGGAGGAA 470
Db 462 GCGAGGATGGGCGCGGAGCGGCGGAGAGGAGCCACAGAGCGGCGCAACCCGAGGAGGAA 521

Oy 471 GCTGCTCCCGACACTGCGCGCGGCTCTGC 501
Db 522 CAGGCGCCAGAGAGCTCCCGCGCGCGCGC 552

RESULT 9
BO931784 867 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8803536 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6426238
DEFINITION 5', mRNA sequence.
ACCESSION BO931784
VERSION BO931784.1 GI:22346815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC <http://mgc.nhl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM2609 row: h column: 11
High quality sequence stop: 627.
Location/Qualifiers
1..867

FEATURES

Location/Qualifiers
1..867

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6426238"
/issue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 101"
/note="Organ: lung; Vector: pOTB7, Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Query Match 6.8%; Score 38.2; DB 5; Length 867;
Best Local Similarity 48.8%; Pred. No. 56;
Matches 103; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Oy 79 TCTGAGGCGAGGCGCAGGCGCAGCTGAGTGCGGAGGCGGCTTGGCTGCTGG 138
Db 70 TCCGCGCCGAGAGGTGTGCTCCAGCTGCGGCGGTCTCACTGCGGCGCAAGGCGCTGGG 129

Oy 139 GCACAGGGGCTGCATACAGCTTACTAGTACATGAGTCCCTGGTGGCCAGCTCTGGA 198
Db 130 CCACCGAGGCCCACTCGGTGCTCCACCCGAGCATGCTTGGGCCCCGACATCAC 189

Oy 199 AGTCTGAGTGAAGCAATGTTCCATTAGAAGATGTGTGCGCGGCGCCATGCCCCCAA 258
Db 190 AGTACGACGAGGAGGACATGTTCTCTTGATGAGGACCCAGGACGCCAGAGCGCATGCG 249

Oy 259 CGTTGCACTCACTGCTTGTGACGGGTTGG 289
Db 250 CATGACAAAGCGCGCGGTGATGATGACAGG 280

RESULT 10
A1970105 393 bp mRNA linear EST 20-OCT-2000
LOCUS w989c02.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2479202 3',
DEFINITION mRNA sequence.
ACCESSION A1970105
VERSION A1970105.1 GI:5766931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 393)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 488 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..393

FEATURES
Location/Qualifiers
1..393
/organism="Homo sapiens"
/mol_type="mRNA"

`/db_xref="taxon:9606"`
`/clone="IMAGE:2479202"`
`/tissue_type="pooled germ cell tumors"`
`/lab_host="DH10B"`
`/clone_lib="NCI CGAP GC6"`
`/note="Vector: pT773D-Pac (Pharmacia) with a modified`
`polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA`
`from the normalized library NCI CGAP GC4 was prepared, and`
`as circles were made in vitro. Following HAP purification,`
`this DNA was used as tracer in a subtractive hybridization`
`reaction. The driver was PCR-amplified cDNAs from a pool`
`of 5,000 clones made from the same library (clonids)`
`1257096-1258631, 1469064-1470983, and 1475592-1476743).`
`Subtraction by Bento Soares and M. Fatima Bonaldo."`

ORIGIN

Query Match	6.7%	Score 38	DB 1	Length 393
Best Local Similarity	54.2%	Pred. No. 54		
Matches	77	Conservative	0	Mismatches 65; Indels 0; Gaps 0;

Qy	Db	Qy	Db
251	150	311	210
CCCCCAAGTTGACAC	CGCCCAAGCTGAGCT	ATCCAGTACAGCCAG	AGGGAAGTCCCA
CTGCTTTGAGGGTGG	CTGAGTCTGAGGGG	CTGCAGAGATCCCT	CAAGACGGGCTG
CTTCCAGCACAGGGT	CGGAGGAGGAGGGG	GCAGTCAATGAACCA	AGGAGGAGG
CTCC	CTCTGCTCC	AGGAGGAGG	AGGAGGAGG
310	209	370	269
371	270		
CTTGGGAACCAATCT	ACTGACATCCCTTGA		
GAGG	GGAGG		
392	291		
CTTGGGAACCAATCT	ACTGACATCCCTTGA		
GAGG	GGAGG		

RESULT	11
AM519131	
LOCUS	AME19131
DEFINITION	ha50f09.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2877161 3', mRNA sequence.

ACCESSION	AM519131	GI:7157213
VERSION	AM519131.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 481)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: image.lnlnl.gov/image/html/tresources.shtml
Seq primer: -40UP from Gibco
High quality sequence start: 419.

FEATURES

```

1..481
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2877161."
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_1b="NCI CGAP Panel"
/note="Organ: pancreas; Vector: pCMV-Sport6; Site 1: SalI
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dn.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

```

Query Match	6.7%	Score 37.8;	DB 1;	Length 481;
Best Local Similarity	56.6%	Pred. No. 64;		
Matches 69;	Conservative	0;	Mismatches 53;	Indels 0;
			Gaps	0;

QY	251	CCCCCAAGTTGACATCTACTGCTTTGACAGGGTGTGGGCTTCAAGTCACAGGGTCCC	310
Db	157	CGCCCAAGCTGGCTCANTTCTGTCCCTGGAGAGAAAGGAGCGAGAGGGCTTCCTGATCCC	216
QY	311	ATCCACGATCCAGCCCAAGTGGCTGTCACAAAGTCCCTCGCACTATGAAACCAAGGAGAG	370
Db	217	AGGAAGTCCCAACCAAGAGCTGTCAACAGGAGCTGGCAAGTCCCTCCAGGAAACA	276

Qy	371	CT	372
Db	277	CT	278

RESULT	12
A1818367	
LOCUS	A1818367
DEFINITION	A1818367 523 bp mRNA linear EST_07-MAR-2000
ACCESSION	wk95jh08.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2419743 3'
VERSION	A1818367
KEYWORDS	A1818367.1 GI:5437446
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE	1 (bases 1 to 523)
AUTHORS	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997)
	Contact: Robert Strusberg, Ph.D.

Email: cgabbs-1@umail.nih.gov
Life Technologies catalog #: 11548-013
DNA sequencing by: Washington University Genome
Clone distribution: NCI-CGAP clone distribution
found through the I.M.A.G.E. Consortium/MLMW at:
www-bio.ljml.gov/bbrrp/image/image.html
Insert Length: 1852 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 428.

FEATURES

```

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2419743"
/tissue_type="adenocarcinoma"
/lab_host="RDH10B"
/clone_1ib="NCI_GCAP_Pan1"
/notes="Orig: pancreas; Vector: pCMV-Sport6; Site 1, SALL1
Site 2: NCI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.72 kb. Life technologies catalog #:
11548-013"

```

ORIGIN

	Query Match	Similarity	6.7%	Score 37.8	DB 1	Length 523
	Best Local	Similarity	56.6%	Pred. No. 65		
	Matches	Conservative	0	Mismatches	53	Indels
					0	Gaps
QY	251	CCCCCAAGTTGCACTCACTGCTTTGAGGGTTGGGGCTTCCACTCAAGGGTCCC	310			
DB	148	CGGCCAAGCTGGCTCACTCTCTGCCCCCTGAGGGGAAGGGGCCAAGGGGCTTCTGTCCC	207			
QY	311	ATCCACGACAGCCAGGTGGCTGACAGAAAGTCCCTGGCACTATGAAACCAAGGAGAG	370			
DB	208	AMGGAATCCACCAACAGAGATCTGGACAACGGGCTGGGCTTGCACTAGTCTCCAGAAACA	267			

OY 371 CT 372
DB 268 CT 269

RESULT 13
LOCUS CF795053/c 592 bp mRNA linear EST 21-OCT-2003
DEFINITION 891143 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CF795053
VERSION CF795053.1 GI:377999626
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae; Suidae;
Sus.

REFERENCE
AUTHORS 1 (bases 1 to 592)
TITLE Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Noneman,D.J., Wray,J.E. and Keele,J.W.
Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: TW8015 row: K column: 23
Seq primer: TAGAAGCAGCTCAGG.
Location/Qualifiers
1..592
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/issue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match 6.7%; Score 37.6; DB 6; Length 592;
Best Local Similarity 50.6%; Pred. No. 75;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
OY 353 TCATGAAACCAAGGAGCTTGGGAAACACATCTGAAGGCGCTTGTATTAGTGA 412
DB 234 TCAGGAAACCAAGGAGAGCTGATATTAATAATACGAGGGGGGCTGTTTGAATATGCA 175
OY 413 GAGGTGGGGCTGGGGCTGGGAGGCCACAGCTGTAGTACAGAGCCAGAGAGCAAGC 472
DB 174 CCGGAGCCAACTGCTCGGGAGGCTGTGAGTGTGTGCTACGGCTGGCCAGCTGTGT 115
OY 473 TGTCTCCAGAGCTGCGCCGCTCTGTGATGAGTCTCTGGCCACTGAGAACAGC 532
DB 114 CCGTCAAGGGCTCGGGCTTCCCTCTTACCTTCTGTGCTCTGTGACACAGGGCTCAGC 55

RESULT 14
LOCUS AA983291 294 bp mRNA linear EST 23-JUN-1998
DEFINITION o656e01.e1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1590360 3',
mRNA sequence.
ACCESSION AA983291
VERSION AA983291.1 GI:3161816
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 294)
TITLE NCI_CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 890 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 273.
Location/Qualifiers
1..294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1590360"
/issue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid5"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AACTGGAAGATTCCGCCCGCCCAATATTTTGTATTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 6.6%; Score 37.2; DB 1; Length 294;
Best Local Similarity 56.6%; Pred. No. 84;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
OY 251 CCCCCAAGCTTGACACTGCTGCTTGGCAGGTTGGGGCTTCCAGTACAGGGTCCC 310
DB 159 CCGGCCATGCTGGTCACTGCTGCTGCCCCCTGAGGAGGAGGCGGCTTCCCGTCCC 218
OY 311 ATCCAGCTACAGCCAGGCTGGCTGCGAAGAGTCCCTCGAGTCATGAAACCAAGGAGG 370
DB 219 AGGGAATCCACACACAGAGTCTGACAGCGGCTGGCGAGTCATCCAGGAAACA 278
OY 371 CT 372
DB 279 CT 280

RESULT 15
LOCUS BG230743 360 bp mRNA linear EST 09-FEB-2001
DEFINITION naF39e01.x1 Soares_NPBMC Homo sapiens cDNA clone IMAGE:4143433 3',
mRNA sequence.
ACCESSION BG230743
VERSION BG230743
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 360)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 JOURNAL Tumor Gene Index
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: M. Bento Soares, Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D. and M. Fatima
 Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNI at:
info@image.llnl.gov
 Seq primer: -40UP from Gibco
 High quality sequence stop: 338.
 Location/Qualifiers

FEATURES
 source
 1..360
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4143433"
 /issue_type="lymphocyte"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Soares NPBC"
 /note="Organ: blood; Vector: pTT73D-Pac; Site 1: NotI;
 Site 2: EcoRI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGCGCCCGGGTTTCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pTT73 vector. Library
 is normalized; constructed in the laboratory of M. Bento
 Soares (University of Iowa)."

ORIGIN

Query Match 6.6%; Score 37.2; DB 2; Length 360;
 Best Local Similarity 56.6%; Pred. No. 87;
 Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```

QY 251 CCCCCACGTTGCACCTGCTTGGACAGGTTGGGGCTTCCAGTCACAGGGTCCC 310
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 153 CCGCCATGCTGCTCAGTTCTGCGCCCTGAGAGGAGGGCCGAGGGCTTCCGTCGCC 212
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 311 ATCCAGTACAGCCCAAGGTGGCTGAGAGAAGTCCCTGCAGTCATGAACCAAGGAGG 370
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 213 AGGGAACCTCCACACAGAGAGTCTGCAGAGGGGCTGGCTGCGAGTCAGTCCAGGAACA 272
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 371 CT 372
    ||
Db 273 CT 274
```

Search completed: March 25, 2006, 18:16:21
 Job time : 3248 secs

Db 17468 CATGACAAAGCGCCGGTGATGATGACGAG 17438

RESULT 2
US-09-949-016-14107
Sequence 14107, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14107
LENGTH: 19319
TYPE: DNA
ORGANISM: Human
US-09-949-016-14107

Query Match
Best Local Similarity 56.6%; Score 37.2; DB 3; Length 19319;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 251 CCCCCCAAGTTGACACTCAGCTTGTGACAGGTTGGCTTCAGTCACAGGGTCCC 310
Db 17937 CCGCCATGCTGGCTCAGTTCTCGCCCTGAGAGGAGGGGCGGCGTCCGTC 17996

Qy 311 ATCCACGTACCAAGCCAGGTGGCTGACAGAGGTCCCTGCAGTCATGAACCAAGGAGG 370
Db 17997 AGGGAAGTCCACACAGGAGTCTGACAGGGGCTGCGCAGTCAGTCCAGGAACA 18056

Qy 371 CT 372
Db 18057 CT 18058

RESULT 3
US-09-621-976-18033
Sequence 18033, Application US/09621976
Patent No. 6639063

GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 18033
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 16
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-18033

Query Match
Best Local Similarity 6.3%; Score 35.4; DB 3; Length 474;
Matches 51; Conservative 147; Mismatches 0; Gaps 0;

Qy 37 CTCACGGTACTCACCCCATGTGGCTGGAGCCGAGGAGCCTCTTAGGACGAGGCCAG 96
Db 4 CTCCTCCYCYGAGCCYCCMARKRSYKGRMYKSMRSGSSCYKSCMCMKRYCSGSYK 63

Qy 97 GAGCCGTCAGTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 156
Db 64 TTTTWTAMMTTGTGKMARRRSGGKTTTMMCSKKKTSCHAGRKXKTYISRMYY 123

Qy 157 GCTTACTGATGACATGAGTCCCTGGTGGCAGCTTGGAGTCTGGAAGTGAAGCAAT 216
Db 124 CYKGCYMMRWKRVSSCCMMYTTGGGSMWTTTMMRRRKSYKRWTKGKKKXTWMA 183

Qy 217 GTTCCCATTAAGGAAGTGTGGCGGCGCATGCCCCCAAGTTGCACTACTGCC 276
Db 184 AMCYTWRSYMMMMRRAAAKTYYYCMMSTKTCMCCMMCCMRBARSCMRSY 243

Qy 277 TTTGACGGTGGGGCTTCCAGTACAGGGGTCACATCCAGTACGACGAGCCGAGTGGC 336
Db 244 TTMCTYYMMYKGRMYMMRGMMKRMYYMMYKKKSNKGSCHMKRAMARKTYTWA 303

Qy 337 AGAAGTCCCTGCGAGTCATGAACCAAGGAGGCTTGGGAACCACTTGAAGGAGAT 396
Db 304 WYTTYRMCYMRKTTYCMWMSYRWMSMTTARAGAMWCMYMYMAARKKTYM 363

Qy 397 GCGTTGATTA 408
Db 364 WAAARGGMTWA 375

RESULT 4
US-09-902-540-6966/C
Sequence 6966, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 6966
LENGTH: 954
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-6966

Query Match
Best Local Similarity 6.2%; Score 35.2; DB 3; Length 954;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 407 TAGTGAGAGGTGGGCTGGCTGGGCAAGGCCACCAAGCTGTGACTGAGCCAGAGCA 466
Db 464 TGTGCGGTGCGCGGCTTTCAGAGGCCCAAGACAGTCTTGCGCGCGCTTGAG 405

Qy 467 GGAAGCTGTCCCAAGCACTGCGCGCGCTTGTGAGTCAAGTCTCTTGCGCACTGAG 526
Db 404 CGGCTCCCTTGACCGGATCCATPACCGAGCGGCATGAGTCCCTCGGCTTCTTCG 345

Qy 527 AACGCTGTAGAGG 542
Db 344 AGCTGCCAGTAGCGAG 329

RESULT 5
US-09-902-540-604
Sequence 604, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 604
LENGTH: 4050
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(4050)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-604

Query Match 6.2%; Score 35.2; DB 3; Length 4050;
Best Local Similarity 53.7%; Pred. No. 5.5;

Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 407 TAGTGAAGAGGTGGGCTGGGCTGGGCAAGGCCACAGGTCTGATGAGGCAAGGCA 466
DB 3587 TCGTCCGGCTCCGCGCTTCCAGAGAGCCCCAGACAGTCTTGGCGCTGCTGGGC 3646
QY 467 GGAAGCTGATCCCGACGACTGCCCGCTCTGAGATGAGTCTCTGGCCAGCTGAG 526
DB 3647 CGGCTCCCTTGAACGGATCATACCGGAGCGGCATGAGTCCCTGGGCTTCTTG 3706
QY 527 AACAGCCTGTAGAGG 542
DB 3707 AGCTGCCAGTAGCGAG 3722

RESULT 6
US-09-843-905A-11

Sequence 11, Application US/09843905A
Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Coaman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 1338
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (513)..(513)
OTHER INFORMATION: unsure
US-09-843-905A-11

Query Match 6.2%; Score 35; DB 3; Length 1338;
Best Local Similarity 49.2%; Pred. No. 4.3;

Matches 92; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 364 AGGAGGCTTGGAAACATCTGAAGGCATGGCTTTGATTTAGTGAAGGCTGGGC 423
DB 794 AGGAGGCTTCTCTCATCACTGTGTGGGCCACATGCTGTGGCCACACCGCGGGGC 853
QY 424 TGGCTGGGCAAGGCACACAGGTCTGATCAGAGCCAGAGGCAAGAGCTGTCCCGAGC 483

DB 854 TGCTGGGGCTCCACACTGAGCACTGAGAGGCCAGCGGACAGAGCAATGACAGGC 913
QY 484 ACTGCGCGCGGCTCTGAGTGAAGTCTCTGCGCCACTGAGAACAGCCTGTAGAGG 543
DB 914 GCGCCAGTCCCGCTGGGCTCAGACTCTGGCTTCCAGCGCCAGCGCGTGGCGCA 973
QY 544 CAGTGC 550
DB 974 CAGCGCC 980

RESULT 7
US-09-949-016-13882

Sequence 13882, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13882
LENGTH: 42610
TYPE: DNA
ORGANISM: Human
US-09-949-016-13882

Query Match 6.2%; Score 35; DB 3; Length 42610;
Best Local Similarity 46.8%; Pred. No. 13;

Matches 110; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 253 CCCCAAGTTCACACTCACTGCTTTGACAGGTTGGGGCTTCCAGTCAAGGCTCCAT 312
DB 13589 CCTGACAGCTTGAACCACTGGCTGCTGTGTGATGAGGCCAGTCCAGAGAGCTG 13648
QY 313 CCAGTACAGCCAGGTGCTGCAAGAGTCTTCCAGTCAAGAACCAAGGAGCT 372
DB 13649 TCACCAAGCGGTAGCACACTCTGACAGTGGGCTGTCTTACAGAGCTCTGTG 13708
QY 373 TGGAAACCATCTGAAGGCGATGCTTTGATTAGTGAAGGCTGGGCTGGG 432
DB 13709 GGGCAGAGTGGGTGAGGCCCAAGAAACATGTGACTGAAAGGAGAGAGGCGAG 13768
QY 433 CAAAGCCACAGGTCTGATCAGAGCCAGAGGCAAGGCTGTCCCACTG 487
DB 13769 CACTCCAGAGGGGTGAGCAGAGACAGCGCGGTGGGATGTGGGCACTGTG 13823

RESULT 8
US-09-949-016-14499/c

Sequence 14499, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14499
; LENGTH: 11276
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(11276)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14499
```

```

Query Match          6.0%; Score 34; DB 3; Length 11276;
Best Local Similarity 51.3%; Pred. No. 17;
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```

QY 346 CTCGCAGTCATGAAACCAAGGAGGCTTGGAAACCACTGAAAGGCGATGGCTTTGAT 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7710 CTGTAAAGCAGAGGCGCACAGGGTGTGAGAGAGCGCTCAGACATGAGGGCGGCTGCCAG 7651
QY 406 TTATGAGAGGAGTGGGCTGGGCTGGGCAAGGCCACCAAGTCTGATCAGAGCCAGAGGC 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7650 TCAAGTAGCAGGAGATGGGCTGGGCAAGGCCACCAAGTCTGATCAGAGCCAGAGGC 7591
QY 466 AGGAAGCTGTGTCGCCAGCACTGCCGCCGCTCT 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7590 AGGAACAGCATCTGGAGACCTGACCACTCCCTCT 7557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 9
US-09-621-976-17564
; Sequence 17564, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 17564
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-17564
```

```

Query Match          6.0%; Score 33.8; DB 3; Length 435;
Best Local Similarity 51.7%; Pred. No. 6.6;
Matches 77; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
```

```

QY 228 AGGAAAGTGTGGCGCGGCGATGCCCCCAAGCTTGACACTGCTTTGCAAGGTT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 166 AGGAAGTGTGCGCGGCGCATTTCCCGCAGCGCCCACTTGCGCTTGAGAGGGGA 225
QY 288 GGGGCTTCAGTCACAGGCTCCCATCCAGTACAGCCAGGCTGAGAGGTCCT 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 226 GGAGGCGCCAGGAGTGAAGTGAAGCAACATCAAAACAGATGCAATTAAT 285
QY 348 CGCAGTCATGAACCAAGGAGGCTTGAG 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 286 CTCGATCCCTTAATCTGTGTGACCTGTG 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 10
US-09-949-016-15483/C
; Sequence 15483, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/231,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15483
; LENGTH: 6872
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(6872)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15483
```

```

Query Match          6.0%; Score 33.8; DB 3; Length 6872;
Best Local Similarity 54.4%; Pred. No. 16;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```

```

QY 390 AGGCATGCTTTATTATTAGTAGAGGAGTGGGCTGGGCTGGGAGGCCACCAAGTCTG 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1625 AGGATGGCTGATTAATTAAGTGAAGGCGTCCAGTCCCAAGAGGCTCTGCGCAAG 1566
QY 450 AGTGAAGCCAGAGGAGGAAAGTGTGCCAGCACTGCCGCCGCTTGCATGCACT 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1565 AACCAGGCAAGGTCAGCTTGGCGCACCTTACCCTGCTCTTGAAGGAGTGTGT 1506
QY 510 CTCTCC 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1505 CTCTCC 1501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 11
US-09-949-016-15980/C
; Sequence 15980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15980
; LENGTH: 239527
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(239527)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15980
```

```

Query Match          6.0%; Score 33.8; DB 3; Length 239527;
Best Local Similarity 48.7%; Pred. No. 52;
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
```



```
QY 336 CAGGTGGCTGCAGAAAGTCCCTCGCATGTAACCAAGGAGGCTTGGAAACCAAT 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194917 CAGGAGACTACTTAAGTGCAGACACAGAGCATGATGCTGTGAGCCAAAC 194858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 CTGAAGGCGATGGCTTTGATTGATGAGAGGGTGGGCTGGGCGCAAGCCACAG 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194857 AGACACTCAACCAAGGCTCGGGTGAAGCTGTGTGGGATCTCTGTCTGCG 194798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446 TCTGAGTCAGACGACGAGGAGGAGGCTGCTCCAGACATGCGCGCGCTTGGGATG 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194797 CTCCACTGGAGAGCATGCTGGCAGCTTGAACAGGCACTCTGAGACCTGCGCTATG 194738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 CAGTCCCTCC 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194737 CAGTTCTTC 194729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-949-016-94858
; Sequence 94858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 94858
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-94858

Query Match
Best Local Similarity 50.6%; Score 33.6; DB 3; Length 601;
Matches 78; Conservative 1; Mismatches 75; Indels 0; Gaps 0;

QY 346 CTCGAGTCATGAAACCAAGGAGGCTTGGAAACCAATCTGAAGGCGATGCTTTGAT 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 149 CTGGTAAGCCAGGCGCCACAGGCTGTGTGAGAGACGGCTCAGACATGGGCGGGCGCCAG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 406 TTAGTGAAGAGGGTGGGCTGGGCTGGGCGCAAGGCCACCAAGTCTGAGTGAAGCCAGAGGC 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 TGAGTAGACAGGAGTGGGCTGGGCGAGGCTGGGGACCCCTGTGGGCTCTCCATGAGACAGA 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 AGGAAGCTGTCCTCCAGCAGCTGCCGCGCTCT 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 AGGAACAGCATCTGGAACCTGACCACTCTCT 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-949-016-12065
; Sequence 12065, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12065
; LENGTH: 108440
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12065

Query Match
Best Local Similarity 6.0%; Score 33.6; DB 3; Length 108440;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 66 AGCGAGGAGACCTCTTCAGGACGAGGCGGCGGCGGCTCAGGGTGTGACGAGGGGT 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81526 AGTGTGTGAACCACTATGGGCTTGGCCAAATCAGCAAGAGGGGCTGGAGAGGGCA 81585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 126 CTGCGCATGTGGGCGACAGGGGCTGCATPACGCTTACTCACTGA 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81586 GCTGTCTTGGGGTACACTGGGACTGGAAGAGTGCCTTCTGA 81629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-949-016-14090
; Sequence 14090, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14090
; LENGTH: 108441
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14090

Query Match
Best Local Similarity 6.0%; Score 33.6; DB 3; Length 108441;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 66 AGCGAGGAGACCTCTTCAGGACGAGGCGGCGGCGGCTCAGGGTGTGACGAGGGGT 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81526 AGTGTGTGAACCACTATGGGCTTGGCCAAATCAGCAAGAGGGGCTGGAGAGGGCA 81585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 126 CTGCGCATGTGGGCGACAGGGGCTGCATPACGCTTACTCACTGA 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81586 GCTGTCTTGGGGTACACTGGGACTGGAAGAGTGCCTTCTGA 81629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-10-104-047-1958
; Sequence 1958, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
```

; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1958
; LENGTH: 2508
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1958

Query Match 5.9%; Score 33.4; DB 3; Length 2508;
Best Local Similarity 48.7%; Pred. No. 15;
Matches 91; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy	364	AGGAGGCTTGGAAACCAATCTGAAGGCGATGCTTTGATTAGTGAGAGGTTGGGC	423
Db	797	AGGACGGCTCTCTCATCGACTGTGTGGGCGCACACTGCTGTGGCGCACACCGCGGGGC	856
Qy	424	TGGGCTGGGCAAGGCCACCAAGTCTGAGTCAGAGCCAGAGCAGAAAGCTGGTCCCAAGC	483
Db	857	TGCTGGGGCTCTCCACACTGAAGCACTGAGGCCCGAGCGCAAGGCAAAATGCAGCGC	916
Qy	484	ACTGCCCGCGCTCTGCGATGCAAGTCTCTTGGCCACCTGAGAAACAGCCTGTAGAGAGG	543
Db	917	GGCCCAAGTGGCCCGGTGGGCTTACGACTCTGGCTTCCCAAGCCAGCCCGTGGCCGCA	976
Qy	544	CAGTGGC	550
Db	977	CAGCGCC	983

Search completed: March 25, 2006, 15:01:59
Job time : 202.5 secs

Db	227	GGGGTCTTGCCATGCTGGGSCAAGGGGCTGCATACGCTTACTGATGACATTCAGTCC	286
Qy	181	CTGGTCCAGCCTCTGGAAGTCTGGAAGTGAACAATGTTTCCATTAAAGAAATGTGTG	240
Db	287	CTGTGTCCAGCCTCTGGAAGTCTGGAAGTGAACAATGTTTCCATTAAAGAAATGTGTG	346
Qy	241	GCCGGCCATGCCCCCAAGCTTGACACTCACTGCTCTTTGACAGGTTGGGGCTTCCAGTC	300
Db	347	GCCGGCCATGCCCCCAAGCTTGACACTCACTGCTCTTTGACAGGTTGGGGCTTCCAGTC	406
Qy	301	ACAGGGGCCATTCACATGACCAAGCCAGGTGCTGCAAGAGTCCCTGCAGTCATGAAA	360
Db	407	ACAGGGGCCATTCACATGACCAAGCCAGGTGCTGCAAGAGTCCCTGCAGTCATGAAA	466
Qy	361	CCAAAGGAGGCTTGGGAAACCAATCTGAAGGGCATGCTTTGATTTAGTGAAGAGGTGG	420
Db	467	CCAAAGGAGGCTTGGGAAACCAATCTGAAGGGCATGCTTTGATTTAGTGAAGAGGTGG	526
Qy	421	GGCTGGGCTGGGGCAAGGCCACCAAGTCTGAGTCAAGGCCCAAGGCCAAGAACTGTCCC	480
Db	527	GGCTGGGCTGGGGCAAGGCCACCAAGTCTGAGTCAAGGCCCAAGGCCAAGAACTGTCCC	586
Qy	481	AGCACTGCCCGCGGCTCTTGCGATGCAGTCTCTCTGGCCACTGAGAAACAGCTGTAGAG	540
Db	587	AGCACTGCCCGCGGCTCTTGCGATGCAGTCTCTCTGGCCACTGAGAAACAGCTGTAGAG	646
Qy	541	AGGCAAGTGGGCTTTCGGACTTC	564
Db	647	AGGCAAGTGGGCTTTCGGACTTC	670

```

RESULT 2
US-10-027-632-142127
Sequence 142127, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827, 129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIORITY APPLICATION NUMBER: US 60/218, 006
PRIORITY FILING DATE: 2000-07-12
PRIORITY APPLICATION NUMBER: US 60/198, 676
PRIORITY FILING DATE: 2000-04-20
PRIORITY APPLICATION NUMBER: US 60/193, 483
PRIORITY FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: US 60/185, 218
PRIORITY FILING DATE: 2000-02-24
PRIORITY APPLICATION NUMBER: US 60/167, 363
PRIORITY FILING DATE: 1999-11-23
PRIORITY APPLICATION NUMBER: US 60/156, 358
PRIORITY FILING DATE: 1999-09-28
PRIORITY APPLICATION NUMBER: US 60/146, 002
PRIORITY FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 142127
LENGTH: 700
TYPE: DNA
ORGANISM: Human
US-10-027-632-142127

```

```

Query Match 564; DB 6; length 700;
Best Local Similarity 99.8%; Pred. No. 2.8e-165;
Matches 563; Conservative 0; Indels 0; Gaps 0
QY 1 GAGGCTGCTCAGAAAAAGAGAGAGCAGACACTTCACGCTGTACTCAGCCCATGTTG 60
db 107 GAGGCTGCTCAGAAAAAGAGAGAGCAGACACTTCACGCTGTACTCAGCCCATGTTG 166

```

Oy	6	GCATGAGGAGGAGAGCTCTCTGAAGGCAAGGCGCAAGGAGACCGCTCAGTGGGTGACCGCA	1 20
Db	167	GCATGAGGCAKAGGAGACCTCTCTGAAGGCAAGGCGCAAGGCGCTCAGTGGGTGACCGCA	2 26
Oy	121	GGGGCTTCTCCATGCTGGGCAACAGGGCTGCAATACGCTTACTCATGTGACATCGAGTCC	180
Db	227	GGGGCTTCTCCATGCTGGGCAACAGGGGCTGCAATACGCTTACTCATGTGACATCGAGTCC	286
Oy	181	CTGGTGCAGACCTCTGGAAAGTCTGGAATGAGACATGTTTCCATTAAAGAAAGTGTG	240
Db	287	CTGGTGCAGACCTCTGGAAAGTCTGGAATGAGACATGTTTCCATTAAAGAAAGTGTG	346
Oy	241	GCCGGCATGACCCGCCCAAGTTGCACTCACTGCGCTTTGCAAGGCTTGGGGCTTCCAGTC	300
Db	347	GCCGGCATGACCCGCCCAAGTTGCACTCACTGCGCTTTGCAAGGCTTGGGGCTTCCAGTC	406
Oy	301	ACAGGGTCCCATCAACGTACACAGGCCAGAGTGGCTGACGAAGGTGCTCGCAGTCATGAAA	360
Db	407	ACAGGGTCCCATCAACGTACACAGGCCAGAGTGGCTGACGAAGGTGCTCGCAGTCATGAAA	466
Oy	361	CCAAGGAGAGCTTGGGAAAACAATCTGAAGGGCATGCGCTTTGATTTAAGTAGAGGGTGG	420
Db	467	CCAAGGAGAGCTTGGGAAAACAATCTGAAGGGCATGCGCTTTGATTTAAGTAGAGGGTGG	526
Oy	421	GGTTGGGCTGGGCAAGGCCACCAAGGTCTGAATCAGACGACAGAGGCAGGAAGCTGTGCTCC	480
Db	527	GGTTGGGCTGGGCAAGGCCACCAAGGTCTGAATCAGACGACAGAGGCAGGAAGCTGTGCTCC	586
Oy	481	AGCAGTCGCCGCGGCTCTGCAGATCAGTCTCTCTGGCCACTTAGAACAAGCTGTAGAG	540
Db	587	AGCAGTCGCCGCGGCTCTGCAGATCAGTCTCTCTGGCCACTTAGAACAAGCTGTAGAG	646
Oy	541	AGGCAATGGCGCTTTTCGACATTC	564
Db	647	AGGCAATGGCGCTTTTCGACATTC	670

```

RESULT 3
US-10-798-652-1
; Sequence 1, Application US/10798652
; Publication No. US20040219582A1
; GENERAL INFORMATION:
; APPLICANT: Yongjun Guo
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISM IN THE
; TITLE OF INVENTION: GGF-3 GENE AND METHODS OF USE THEREOF
; FILE REFERENCE: 3982-P03336US01
; CURRENT APPLICATION NUMBER: US/10/798, 652
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/455,698
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: n is cytosine or thymine
US-10-798-652-1

```

[illegible]

QY	121	GGGGCTTGGCAATGGTGGGCA	CAGGGGCTGCATACAGCTTACTCAGTGCAATCGAGTCC	180
Db	121	GGGGCTTGGCAATGGTGGGCA	CAGGGGCTGCATACAGCTTACTCAGTGCAATCGAGTCC	180
QY	181	CTGTGTCCAGCCTCTGGAAGTCTGGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG	240	
Db	181	CTGTGTCCAGCCTCTGGAAGTCTGGAAGTGAAGCAATGTTTCCATTAAAGAAAGTGTG	240	
QY	241	GGCGGCATGCCCCCCCAAGTTGCACTGACTGCTTTGCAAGGTTGTGGGCTTCCAGTC	300	
Db	241	GGCGGCATGCCCCCCCAAGTTGCACTGACTGCTTTGCAAGGTTGTGGGCTTCCAGTC	300	
QY	301	ACAGGGTCCCATCCAGCTACAGCCAGGGGTGTCAGAAAGGCTCTGCAGTCATGA	360	
Db	301	ACAGGGTCCCATCCAGCTACAGCCAGGGGTGTCAGAAAGTCTCTGCAGTCATGA	360	
QY	361	CCAAAGGAGGCTTTGGAAA	CAATCTGAAGGAGCATGAGCTTTGATTTAGTGAAGGGTGG	420
Db	361	CCAAAGGAGGCTTTGGAAA	CAATCTGAAGGAGCATGAGCTTTGATTTAGTGAAGGGTGG	420
QY	421	GGCTGGGCTGGGCAAGGGCCACAGGCTGAAGTCAGAGCCAGAGGCAAGGAAAGCTGATCCCC	480	
Db	421	GGCTGGGCTGGGCAAGGGCCACAGGCTGAAGTCAGAGCCAGAGGCAAGGAAAGCTGATCCCC	480	
QY	481	AGCATGCCCCGCGCTCTGCGATGCAATGCTCTCTGGCCACCTGAGAA	CAGCCTGTAGAG	540
Db	481	AGCATGCCCCGCGCTCTGCGATGCAATGCTCTCTGGCCACCTGAGAA	CAGCCTGTAGAG	540
QY	541	AGGCAATGGCGCTTTTGGACTTC	564	
Db	541	AGGCAATGGCGCTTTTGGACTTC	564	

```

RESULT 4
US-10-322-281-642
; Sequence 642, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945200100
; CURRENT PILING DATE: US/10/322,281
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 642
; LENGTH: 29340
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-642

Query Match          98.0%; Score 553; DB 7; Length 29340;
Best Local Similarity 99.8%; Pred. No. 9.9e-162;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1.

QY      1  GCAGCCCTGCTCAGAAAAACAGAAAGACGACAGACACTCACGGTGTACTCACCCCCCATGTG 60
DB      3566  GCAGCCCTGCTCAGAAAAACAGAAAGACGACAGACACTCACGGTGTACTCACCCCATGTG 3625

QY      61  GCTGAGGCGAGGAGGAGCCTCTCTGAGAGCAGGGGCGAGGCGAAGCCCTGTCAAGTGGGTGACGGCA 120
DB      3626  GCTGAGGCGAGGAGGAGCCTCTCTCTGAGAGCAGGGGCGAGGCGAAGCCCTGTCAAGTGGGTGACGGCA 3685

QY      121  GGGGTCTTGCATGTGTGGGCAACAGGGGCTGTACATACAGCTTACTCAGTGACATCGAGTCC 180
DB      3686  GGGGTCTTGCATGTGTGGGCAACAGGGGCTGTACATACAGCTTACTCAGTGACATCGAGTCC 3745

QY      181  CTGGTGCAGGCTCTGGAAGTCTGGAAGTGAACAATGTTTCCCATTAAGAAAGTGTGTG 240
DB      3746  CTGGTGCAGGCTCTGGAAGTCTGGAAGTGAACAATGTTTCCCATTAAGAAAGTGTGTG 3805

```

QY	241	ACCGGCATGCCCCCAACGTTTGAACATCACTACCTTTGGAGGGTTGGGCTTCCAGTC	300
Db	3806	GCCCCCATGCCCCCAACGTTTGACACTACTCTCTTTGAGGGTTGGGGCTTCCAGTC	3865
QY	301	ACAGGGTCCATCCACGTAACGACGCCCAGGTGGCTGCAAGAGGTCCCTGCAGTCATGAAA	360
Db	3866	ACAGGGTCCATCCACGTAACGACGCCCAGGTGGCTGCAAGAGGTCCCTGCAGTCATGAAA	3925
QY	361	CCAAAGGAGGCTTTGGGAAACCAATCTTGAAGGGGATGGCTTTGATTTAGTGAAGGGTGG	420
Db	3926	CCAAAGGAGGCTTTGGGAAACCAATCTTGAAGGGGATGGCTTTGATTTAGTGAAGGGTGG	3985
QY	421	GGCTGGGCTGGGCAAGGCGCACAGGCTTGAATCAGAGCCAGAGGCAAGAGCTGGTCCCC	480
Db	3986	GGCTGGGCTGGGCAAGGCGCACAGGCTTGAATCAGAGCCAGAGGCAAGAGCTGGTCCCC	4045
QY	481	A-GCACTGCCCCGCGGCTCTTGCGATGCAATCTCTTGCCCACTGAGAACAGCCTGTGAGA	539
Db	4046	AGGCACTGCCCCGCGGCTCTTGCGATGCAATCTCTTGCCCACTGAGAACAGCCTGTGAGA	4105
QY	540	GAGGCAAGTGGGCTCTTTCCGACCTTC	564
Db	4106	GAGGCAAGTGGGCTCTTTCCGACCTTC	4130

```

RESULT 5
US-10-027-632-142126
; Sequence 142126, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Manq, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,353
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142126
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142126

Query Match      98.0%; Score 552.6; DB 5; Length 855;
Best Local Similarity 99.6%; Pred. No. 7,9e-162;
Matches 563; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY      1  GCAACCTCGCTCAAGAAACAGAAAGACGAGACACATCAACGGTAGTCACCCCAATGTG 60
DB      108  GCAACCTCGCTCAAGAAACAGAAAGACGAGACACATCAACGGTAGTCACCCCAATGTG 167
QY      61  GCTGAGAGCGAGGAGGAGCCTCTTGAGGCAAGGAGCCAGGCAACCGTCAGGTGGTGAACGGCA 120
DB      168  GCTGAGAGTGAAGGAGGAGCCTCTTGAGGCAAGGAGCCAGGCAACCGTCAGGTGGTGAACGGCA 227
QY      121  GGGGCTCTTGCCATGATGAGGCAACAGGGGCTGCATACAGCTTACTGATGACATGAGTCC 180
DB      228  GGGGCTCTTGCCATGATGAGGCAACAGGGGCTGCATACAGCTTACTGATGACATGAGTCC 287

```

QY	181	CTGTGTCCAGCCCTCTGGAAGTGTGGAAAGGACATGTTTCCATTAAAGAAAGTGCTG	240
Db	288	CTGTGTCCAGCCTCTGGAAGTGTGGAAAGTGCATATTTCCATTAAAGAAAGTGCTG	347
QY	241	GCCGCGCATGCCCCCACCAGTTGCACTCATCTGCTTTGCAAGGGTTGGGGCTTCCAGTC	300
Db	348	GCCGCGCATGCCCCCACCAGTTGCACTCATCTGCTTTGCAAGGGTTGGGGCTTCCAGTC	407
QY	301	ACAGGGTCCCATCTCACGTACCAAGCCAGTGGCTTGCAGAAAGTCCCTTCGACATGAAA	360
Db	408	ACAGGGTCCCATCTCACGTACCAAGCCAGTGGCTTGCAGAAAGTCCCTTCGACATGAAA	467
QY	361	CCAAAGGAGGCTTTGGGAAACCACTCTGAAGGGCATGGCTTTGATTTAGTGAAGAGGGTGG	420
Db	468	CCAAAGGAGGCTTTGGGAAACCACTCTGAAGGGCATGGCTTTGATTTAGTGAAGAGGGTGG	527
QY	421	GGCTGGGGCTGGGCAAGGCCACCAAGGCTTGAGTCAGAGCCAAAGGACAGGAAGTGGTCCC	480
Db	528	GGCTGGGGCTGGGCAAGGCCACCAAGGCTTGAGTCAGAGCCAAAGGACAGGAAGTGGTCCC	587
QY	481	A-GCAGTGGCCCGCGCTCTTGCGATGCACTCTCTGSCACCTTGAGAAAGCCTGTAGA	539
Db	588	AGGCACTGCCCCCGCTCTTGCGATGCACTCTCTGSCACCTTGAGAAAGCCTGTAGA	647
QY	540	GAGGCACTGGCGTCTTTGGAGCTTC	564
Db	648	GAGGCACTGGCGTCTTTGGAGCTTC	672

```

RESULT 6
US-10-027-632-142126
Sequence 142126, Application US/10027632
Publication NO. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027, 632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218, 006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198, 676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193, 483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185, 218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167, 363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156, 358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146, 002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO. 142126
LENGTH: 855
TYPE: DNA
ORGANISM: Human
US-10-027-632-142126

```

	Query Match	Similarity	Score	DB	Length	
	Best local	99.6%	552.6	162	855	
	Matches	563	Conservative	1	Mismatches	0
					Indels	1
					Gaps	1
Qy	1	GCAGCCTCGCCTCAGAAAAACAGAAAGACGACACACTCACGGTGTACTTACCCCATGTG	60			
Pb	108	GCAGCCTCGCCTCAGAAAAACAGAAAGACGACACACTCACGGTGTACTTACCCCATGTG	167			
Oy	61	GCTGAGGCGAAGGAGCCTCTCTGAGGACGAGGACGAGGCTCAGTGGGTGACGCA	120			

Db	166	GCCTGAGAGTGAAGGAGCCCTCTGAGGCGAGGGCCAGAGGACGGCTGACGTGGGTGACGGCA	227
Oy	121	GGGGTCTTGGCCATGCTGGGGCACAGGGGCTGCAATACGTTACTAGTGAACATTCGAGTCC	180
Db	228	GGGGTCTTGGCCATGCTGGGGCACAGGGGCTGCAATACGTTACTAGTGAACATTCGAGTCC	287
Oy	181	CTGGTCCAGCCTCTGGAAGTCTGGAGTGAAGCATGTTCCATTAGGAAGTGTGTG	240
Db	288	CTGGTCCAGCCTCTGGAAGTCTGGAGTGAAGCATGTTCCATTAGGAAGTGTGTG	347
Oy	241	GCCGCGCATGCCCCCCCAACGTTGCACATCACTGCTTTTGCAAGGTTGGGGCTTCCAGTC	300
Db	348	GCCGCGCATAGCCCCCCCAAGTTGCACATCACTGCTTTTGCAAGGTTGGGGCTTCCAGTC	407
Oy	301	ACAGGGTCCCATCCAGTACACAGCCCAAGTGGCTGCAAGAGTCCCTCGCAGTCATGATAA	360
Db	408	ACAGGGTCCCATCCAGTACACAGCCCAAGTGGCTGCAAGAGTCCCTCGCAGTCATGATAA	467
Oy	361	CCAAGGAGAGGCTTGGGAAAACAATCTGAAGGGCATGGCTTTGATTTAGTGAAGGGTGG	420
Db	468	CCAAGGAGAGGCTTGGGAAAACAATCTGAAGGGCATGGCTTTGATTTAGTGAAGGGTGG	527
Oy	421	GGCTGGGCTGGGGCAAGGCCACAGGCTCTGAATCTCAAGCCAGAGGCAGAAAGCTGTCCC	480
Db	528	GGCTGGGCTGGGGCAAGGCCACAGGCTCTGAATCTCAAGCCAGAGGCAGAAAGCTGTCCC	587
Oy	481	A-GCATTGCGCGCGCTCTGCGATGAGTCTTCCTGGCCAACCTGAGAAAGCCTGTAGA	539
Db	588	AGGCACTGCGCGCGCTCTGCGATGAGTCTTCCTGGCCAACCTGAGAAAGCCTGTAGA	647
Oy	540	GAGGCAGTGGGCTTTTGGGACTTC	564
Db	648	GAGGCAGTGGGCTTTTGGGACTTC	672

```

RESULT 7
US-09-925-065A-734471
; Sequence 734471, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 734471
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-734471

```

Query Match	7.1%	Score 39.8	DB 4	Length 541
Best Local Similarity	64.8%	Pred. No. 0.056		
Matches	59	Conservative	0	Mismatches 32; Indels 0; Gaps 0
OY	4	GCCTTGCCTCGAAGAAAGAGCGACGACACTCAGCGTGAATCCACCCCATGTGCT	63	
Db	442	GCCTTCCCTCGGAACCGAAGGCTGGGACCCCTCGTGGGCTCTACCCCTCATTTGAA	501	
OY	64	GGAGCGAAGGAGGCTCTCTGAGGCGAGGCCA	94	

;/ Prior Application removed - See File Wrapper or Palm
;/ NUMBER OF SEQ ID NOS: 550
;/ SEQ ID NO 10
;/ LENGTH: 594
;/ TYPE: PRT
;/ ORGANISM: Homo Sapien
US-10-146-731-10

Query Match 6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.24; Mismatches 277; Indels 0; Gaps 0;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

QY 23 AAGACGACGACATCTACGAGTACACCCCATGCTGAGCGAGGAGGAGCTCT 82
DB 56 VALAAESPVNMMPMWGDITNMIDRFVRAHLHDHPYTPPLTTTISPEQESDERKCYNER 115
QY 83 GAGGACGAGGCGACGAGCGCTCAGGTGAGTACGCGAGGAGGCTTCCATGCTGGGAC 142
DB 116 YRGLVQNDPAGISEEQCLYQIYIDELVGLQRPSEDEKKKLAEKASIGTYEDSTVAEV 175
QY 143 AAGGCTCATACAGCTTACTCAGTACGACATCGAGTCCCTGGTGCAGCGCTTGAAGTC 202
DB 176 EKAAREKPEEBSAAEESNSDEDEVIPDIDVEVDVDELNOBQVADLNKQATTYGMADDF 235
QY 203 TGAAGTGAAGCAATGTTTCCATTAGAAGAGTGTGTGCGCGCCATGCCCCCAACGTT 262
DB 236 VMLRKDXEABAIKHAALKEEKAMYSGRRSRQRREFREKRLRGKTISSPSYARDBP 295
QY 263 GCACATCTACTGCTTTCAGAGGTTGGGCTTCCAGTACAGGTCCTCCACTACCA 322
DB 296 TYDPYKRSPESSSSRSRSPPTGPREKITYTITSFGSDEBMAAAAAAASGVTTGK 355
QY 323 GCCAAGTGGCTGCAAGAGTCCCTGCGACATCAAGAACCAAGGAGGCTTGGAAACCA 382
DB 356 PPAPPOGPPAPGRNARSARRSSSSSSASRTSSRSRSGGTYYSGRHA 415
QY 383 CATCTGAAGGCGATGCTTTGATTAGTAGAGAGGTGGGCTGCGCTGGGCAAGCCACC 442
DB 416 RSRSRWSRSRSTRSRYSRGRHSGGSRDGRHRSPPARRGGYPRRRSRSHS 475
QY 443 AGCTGAGTCAAGAGCGCAAGAGCAAGTGTCTCCCGACACTGCCCCCCTCTGG 502
DB 476 GDRYRGGGRLRHSSSRSSWSLSPSRSLTTRSHSPSPGSRSRSRSGSPSPS 535
QY 503 ATGCACT 509
DB 536 PAREKLT 542

RESULT 11
US-10-140-472-10

;/ Sequence 10, Application US/10140472
;/ Publication No. US2003013888A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Baker, Kevin P.
;/ APPLICANT: Beresini, Maureen
;/ APPLICANT: Deforge, Laura
;/ APPLICANT: Desnoyers, Luc
;/ APPLICANT: Filvaroff, Ellen
;/ APPLICANT: Gao, Wei-Qiang
;/ APPLICANT: Geritsen, Mary E.
;/ APPLICANT: Goddard, Audrey
;/ APPLICANT: Godowski, Paul J.
;/ APPLICANT: Gurney, Austin L.
;/ APPLICANT: Sherwood, Steven
;/ APPLICANT: Smith, Victoria
;/ APPLICANT: Stewart, Timothy A.
;/ APPLICANT: Tumas, Daniel
;/ APPLICANT: Watanabe, Colin K
;/ APPLICANT: Wood, William
;/ APPLICANT: Zhang, Zemin
;/ APPLICANT: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;/ TITLE OF INVENTION: ACIDS ENCODING THE SAME

;/ FILE REFERENCE: P3330R1C168
;/ CURRENT APPLICATION NUMBER: US/10/140,472
;/ CURRENT FILING DATE: 2002-05-06
;/ Prior Application removed - See File Wrapper or Palm
;/ NUMBER OF SEQ ID NOS: 550
;/ SEQ ID NO 10
;/ LENGTH: 594
;/ TYPE: PRT
;/ ORGANISM: Homo Sapien
US-10-140-472-10

Query Match 6.7%; Score 37.8; DB 6; Length 594;
Best Local Similarity 7.0%; Pred. No. 0.24; Mismatches 277; Indels 0; Gaps 0;
Matches 34; Conservative 176; Mismatches 277; Indels 0; Gaps 0;

QY 23 AAGACGACGACATCTACGAGTACACCCCATGCTGAGCGAGGAGGAGCTCT 82
DB 56 VALAAESPVNMMPMWGDITNMIDRFVRAHLHDHPYTPPLTTTISPEQESDERKCYNER 115
QY 83 GAGGACGAGGCGACGAGCGCTCAGGTGAGTACGCGAGGAGGCTTCCATGCTGGGAC 142
DB 116 YRGLVQNDPAGISEEQCLYQIYIDELVGLQRPSEDEKKKLAEKASIGTYEDSTVAEV 175
QY 143 AAGGCTCATACAGCTTACTCAGTACGACATCGAGTCCCTGGTGCAGCGCTTGAAGTC 202
DB 176 EKAAREKPEEBSAAEESNSDEDEVIPDIDVEVDVDELNOBQVADLNKQATTYGMADDF 235
QY 203 TGAAGTGAAGCAATGTTTCCATTAGAAGAGTGTGTGCGCGCCATGCCCCCAACGTT 262
DB 236 VMLRKDXEABAIKHAALKEEKAMYSGRRSRQRREFREKRLRGKTISSPSYARDBP 295
QY 263 GCACATCTACTGCTTTCAGAGGTTGGGCTTCCAGTACAGGTCCTCCACTACCA 322
DB 296 TYDPYKRSPESSSSRSRSPPTGPREKITYTITSFGSDEBMAAAAAAASGVTTGK 355
QY 323 GCCAAGTGGCTGCAAGAGTCCCTGCGACATCAAGAACCAAGGAGGCTTGGAAACCA 382
DB 356 PPAPPOGPPAPGRNARSARRSSSSSSASRTSSRSRSGGTYYSGRHA 415
QY 383 CATCTGAAGGCGATGCTTTGATTAGTAGAGAGGTGGGCTGCGCTGGGCAAGCCACC 442
DB 416 RSRSRWSRSRSTRSRYSRGRHSGGSRDGRHRSPPARRGGYPRRRSRSHS 475
QY 443 AGCTGAGTCAAGAGCGCAAGAGCAAGTGTCTCCCGACACTGCCCCCCTCTGG 502
DB 476 GDRYRGGGRLRHSSSRSSWSLSPSRSLTTRSHSPSPGSRSRSRSGSPSPS 535
QY 503 ATGCACT 509
DB 536 PAREKLT 542

RESULT 12
US-10-141-761-10

;/ Sequence 10, Application US/10141761
;/ Publication No. US20030148432A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Baker, Kevin P.
;/ APPLICANT: Beresini, Maureen
;/ APPLICANT: Deforge, Laura
;/ APPLICANT: Desnoyers, Luc
;/ APPLICANT: Filvaroff, Ellen
;/ APPLICANT: Gao, Wei-Qiang
;/ APPLICANT: Geritsen, Mary E.
;/ APPLICANT: Goddard, Audrey
;/ APPLICANT: Godowski, Paul J.
;/ APPLICANT: Gurney, Austin L.
;/ APPLICANT: Sherwood, Steven
;/ APPLICANT: Smith, Victoria
;/ APPLICANT: Stewart, Timothy A.
;/ APPLICANT: Tumas, Daniel
;/ APPLICANT: Watanabe, Colin K
;/ APPLICANT: Wood, William

RESULT 13
 US-10-142-885-10
 Sequence 10, Application US/10142885
 Publication No. US20030157604A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Beresini, Maureen
 APPLICANT: DeForge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Flvavoff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Andrew
 APPLICANT: Gowerski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.

RESULT 14
US-10-158-790-10
Sequence 10, Application US/10158790
Publication No. US20030180879A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Berselli, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlisen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.

```
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C448
CURRENT APPLICATION NUMBER: US/10/158,790
CURRENT FILING DATE: 2002-05-30
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRT
ORGANISM: Homo Sapien
US-10-158-790-10
```

Query Match 6.7%, Score 37.8, DB 6, Length 594;
Best Local Similarity 7.0%, Pred. No. 0.24, Mismatches 277, Indels 0, Gaps 0;
Matches 34, Conservative 176, Mismatches 277, Indels 0, Gaps 0;

```
QY 23 AAGACGACGACACTCAGTCACTGAGTCACTCCCATGCTGAGAGCGAGGAGCTCCT 82
DB 56 VALAABSPVNMWPMQGDNNMIDRFDVRAHLDHIDPYPPLTTTISPEQESDERKCNVER 115
QY 83 GAGGACGAGGCGGAGGAGCGGCTCAGTGGTGAACGAGGAGGCTTGGCATGTGGGAC 142
DB 116 YRGLVQNDPAGISEQCILYIYIDELVGLQRPSEDEKKLAEKASIGTYEDSTVAEV 175
QY 143 AGGGCTGCATACAGCTTACTGATGACATGCAAGTCCCTGGTGCAGGCTTGGAGATC 202
DB 176 EKAABKEPEEBSAAEBSNSDEVDIPIDIVENVDELQBVADLNQATTYGMADDF 235
QY 203 TGAAGTGAAGCATGTTTCCCATTAAGAAAGTGTGTGGCCGACATGCCCCCAACGTT 262
DB 226 VMLRKDEEBAIKHAKALEEKAMYSGRSRQRFREKRLRGKISPSYARRDSP 295
QY 263 GCACACTCAGTCCCTTGGAGGCTTGGGCTTCCAGTCAAGGATCCCATCCAGTACCA 322
DB 296 TYDPYKSPSSSSSSRSRSPTPGREKITFTFSFGSDDEMAAAAAAASGVTTGK 355
QY 322 GCCCAGGTGGCTGCAGAAAGTCCCTCGCAGTCAAGAACCAAGGAGGCTTGGAAACA 382
DB 356 PPAPPOPGPAPGRNASARRSSSSSSASRTSSRSRSGGYYRSGRHA 415
QY 383 CATCTGAAGGCGATGCTTTGATTAGTAGAGAGGTGGGCTGGCTGGCAAGCCACC 442
DB 416 RSRSRWSRSRSGRNRYSRGRHSGGSRDGRYRSPARRGGYGPARRSRSHS 475
QY 443 AGGTCTGATCAGAGCCAGAGGAGAGAGTGGTCCCGACACTGCCCCCGCTTGGG 502
DB 476 GDRYRGGGRLHHSSRSRWSLSPSRSRLTFRSHSPSPQSRSRSGSPSPS 535
QY 503 ATGCACT 509
DB 536 PAREKIT 542
```

```
RESULT 15
US-10-137-871-10
Sequence 10, Application US/10137871
Publication No. US20030207350A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltsen, Mary E.
```

```
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C153
CURRENT APPLICATION NUMBER: US/10/137,871
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-871-10
```

Query Match 6.7%, Score 37.8, DB 6, Length 594;
Best Local Similarity 7.0%, Pred. No. 0.24, Mismatches 277, Indels 0, Gaps 0;
Matches 34, Conservative 176, Mismatches 277, Indels 0, Gaps 0;

```
QY 23 AAGACGACGACACTCAGTCACTGAGTCACTCCCATGCTGAGAGCGAGGAGCTCCT 82
DB 56 VALAABSPVNMWPMQGDNNMIDRFDVRAHLDHIDPYPPLTTTISPEQESDERKCNVER 115
QY 83 GAGGACGAGGCGGAGGAGCGGCTCAGTGGTGAACGAGGAGGCTTGGCATGTGGGAC 142
DB 116 YRGLVQNDPAGISEQCILYIYIDELVGLQRPSEDEKKLAEKASIGTYEDSTVAEV 175
QY 143 AGGGCTGCATACAGCTTACTGATGACATGCAAGTCCCTGGTGCAGGCTTGGAGATC 202
DB 176 EKAABKEPEEBSAAEBSNSDEVDIPIDIVENVDELQBVADLNQATTYGMADDF 235
QY 203 TGAAGTGAAGCATGTTTCCCATTAAGAAAGTGTGTGGCCGACATGCCCCCAACGTT 262
DB 226 VMLRKDEEBAIKHAKALEEKAMYSGRSRQRFREKRLRGKISPSYARRDSP 295
QY 263 GCACACTCAGTCCCTTGGAGGCTTGGGCTTCCAGTCAAGGATCCCATCCAGTACCA 322
DB 296 TYDPYKSPSSSSSSRSRSPTPGREKITFTFSFGSDDEMAAAAAAASGVTTGK 355
QY 322 GCCCAGGTGGCTGCAGAAAGTCCCTCGCAGTCAAGAACCAAGGAGGCTTGGAAACA 382
DB 356 PPAPPOPGPAPGRNASARRSSSSSSASRTSSRSRSGGYYRSGRHA 415
QY 383 CATCTGAAGGCGATGCTTTGATTAGTAGAGAGGTGGGCTGGCTGGCAAGCCACC 442
DB 416 RSRSRWSRSRSGRNRYSRGRHSGGSRDGRYRSPARRGGYGPARRSRSHS 475
QY 443 AGGTCTGATCAGAGCCAGAGGAGAGAGTGGTCCCGACACTGCCCCCGCTTGGG 502
DB 476 GDRYRGGGRLHHSSRSRWSLSPSRSRLTFRSHSPSPQSRSRSGSPSPS 535
QY 503 ATGCACT 509
DB 536 PAREKIT 542
```

Search completed: March 25, 2006, 18:14:34
odb time : 682.5 secs